

La recombinaison illégitime dans les cellules de mammifère

La recombinaison illégitime (ou non homologue) est un processus cellulaire conduisant à la réunion de deux fragments d'ADN ne présentant pas ou très peu d'homologie de séquence. Elle est à l'origine de diverses anomalies du génome (délétion, insertion, translocation, amplification génique, intégration virale) impliquées dans l'apparition de maladies génétiques ou le développement tumoral. Elle permet l'élimination des cassures double-brin accidentelles de la chaîne nucléotidique, consécutives à des erreurs du métabolisme normal de l'ADN (réplication, transcription, réparation) ou à la présence de lésions sur l'ADN produites par les agents génotoxiques. Certains motifs (séquences répétées de type *Alu*, palindromes, séquences alternées de purines-pyrimidines, site de clivage des topo-isomérases...) seraient des points chauds de recombinaison illégitime, régions fragiles ou séquences régulatrices plus accessibles aux nucléases capables d'introduire des cassures dans l'ADN.

Anne Stary

La recombinaison génétique désigne n'importe quel processus au cours duquel l'interaction d'au moins deux éléments différents d'acide nucléique a pour conséquence un changement dans la liaison physique de gènes ou de parties de gènes. La nature des séquences qui interagissent permet de distinguer trois types de recombinaison : (1) la recombinaison générale (ou homologue) qui nécessite une identité importante entre les séquences qui interagissent ; (2) la recombinaison spécialisée qui intéresse des séquences parti-

culières et des protéines spécifiques. Une courte région d'homologie (quelques paires de bases) à l'intérieur des séquences est nécessaire, mais non suffisante, pour que se produise la recombinaison. Dans les cellules de mammifère, la recombinaison spécialisée est impliquée dans la maturation des immunoglobulines, pendant la différenciation des lymphocytes ; (3) la recombinaison illégitime (ou non homologue) qui est définie par la recombinaison entre deux séquences d'ADN ne présentant pas ou très peu d'homologie de séquence [1-3]. L'identification

ADRESSE

A. Stary : chargée de recherche au Cnrs. Laboratoire de génétique moléculaire, Institut de recherches sur le cancer, 7, rue Guy-Moquet, 94800 Villejuif, France.

RECOMBINAISON **HOMOLOGUE** : au moins 14 nucléotides consécutifs communs entre les deux séquences parentales (P1 et P2) impliquées au site de recombinaison isolé dans la séquence réarrangée (R).

A

P1	G	<u>C</u>	<u>T</u>	<u>C</u>	<u>A</u>	<u>T</u>	<u>T</u>	<u>G</u>	<u>C</u>	<u>A</u>	<u>G</u>	<u>T</u>	<u>A</u>	<u>C</u>	<u>C</u>	<u>G</u>	<u>G</u>	<u>T</u>	<u>A</u>	<u>T</u>	<u>A</u>	<u>T</u>	<u>G</u>	<u>G</u>	<u>A</u>	<u>C</u>	<u>G</u>	<u>T</u>	<u>T</u>			
	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	
R	G	<u>C</u>	<u>T</u>	<u>C</u>	<u>A</u>	<u>T</u>	<u>T</u>	<u>G</u>	<u>C</u>	<u>A</u>	<u>G</u>	<u>T</u>	<u>A</u>	<u>C</u>	<u>C</u>	<u>G</u>	<u>G</u>	<u>T</u>	<u>A</u>	<u>A</u>	<u>C</u>	<u>T</u>	<u>C</u>	<u>G</u>	<u>G</u>	<u>A</u>	<u>T</u>	<u>C</u>	<u>C</u>	<u>G</u>		
	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	
P2	T	G	T	A	A	G	T	G	T	A	G	T	A	A	C	T	C	G	G	T	A	A	C	T	C	G	G	A	T	C	C	G

RECOMBINAISON **NON-HOMOLOGUE** : pas ou peu de nucléotides consécutifs communs entre les deux séquences parentales (P1 et P2) impliquées au site de recombinaison isolé dans la séquence réarrangée (R).

B

P1	T	T	A	G	T	C	A	T	T	G	C	A	G	T	G	G	T	A	A	T	C	G	T	A	G	C	C	A	A	G	C	G		
	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•		
R	T	T	A	G	T	C	A	T	T	G	C	A	G	T	G	G	T	A	A	C	T	C	G	T	A	A	C	T	T	G	G	A	T	
	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	
P2	G	A	C	T	G	T	A	A	G	C	G	T	A	A	T	G	G	T	A	A	C	T	C	G	G	T	A	A	C	T	T	G	G	A

C

P1	T	T	A	G	T	C	A	T	T	G	C	A	G	T	G	G	T	A	A	T	G	C	C	A	G	C	C	A	T	G	C	G		
	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•		
R	T	T	A	G	T	C	A	T	T	G	C	A	G	T	G	G	C	C	G	G	T	A	A	C	T	T	G	G	A	T	G	G	A	T
	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
P2	G	A	A	T	G	T	C	T	A	T	G	T	A	A	C	C	T	C	G	C	C	G	G	T	A	A	C	T	T	G	G	A	T	G

D

P1	T	A	G	T	C	A	T	T	G	C	A	G	T	G		C	A	C	T	C	C	T	C	G	C	C	A	T	G	C	G		
	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•		
R	T	A	G	T	C	A	T	T	G	C	A	G	T	G	A	T	C	A	C	C	G	G	T	A	A	C	T	T	G	G	A	T	G
	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
P2	A	C	T	G	T	A	A	G	C	G	T	A	G	T	T	C	A	C	C	G	G	T	A	A	C	T	T	G	G	A	T	G	G

Figure 1. Exemples de jonctions de recombinaison homologue et non homologue. Les points indiquent les nucléotides communs aux deux séquences impliquées dans la recombinaison. Les bases communes à la jonction de recombinaison sont soulignées et les nucléotides supplémentaires à cette jonction sont en gras. Les nucléotides en rouge correspondent aux nucléotides communs aux deux séquences parentales.

moléculaire [4] et à vingt-cinq pour la recombinaison intermoléculaire [5] (figure 1A). La définition de la recombinaison illégitime pourrait donc s'appliquer à tout phénomène de réarrangement entre des séquences présentant moins de quatorze nucléotides consécutifs identiques au site de recombinaison ou dans les séquences proches de ce site (figures 1B, 1C, 1D). Cependant, dans les cellules de mammifère, toutes les jonctions de recombinaison non homologue, répertoriées en tant que telles dans la littérature, ne font intervenir au plus que sept nucléotides consécutifs d'identité [1]. Dans les cellules de mammifère, en utilisant des modèles d'étude permettant de différencier les réarrangements de type homologue ou non homologue, il apparaît que la recombinaison illégitime se produit avec une fréquence plus élevée que la recombinaison homologue. Ainsi, le principal mode d'intégration dans le génome d'une séquence exogène et homologue à un gène cellulaire n'est pas le remplacement du gène endogène par recombinaison homologue (recombinaison ciblée) mais l'intégration au hasard de la séquence par recombinaison illégitime et ce, avec une fréquence cent à mille fois supérieure [2]. La recombinaison illégitime est également la règle générale dans des systèmes de recombinaison extrachromosomique lorsque les conditions expérimentales permettent de différencier les deux types de recombinaison [6]. La recombinaison illégitime est donc un processus cellulaire important dans les cellules de mammifère, qui peut être à l'origine de diverses anomalies du génome entraînant l'apparition de maladies génétiques ou le développement tumoral. De plus,

des phénomènes de recombinaison nécessite donc la détermination du degré d'homologie entre les deux séquences d'ADN impliquées au site de recombinaison (figure 1). Trois types de jonctions de recombinaison illégitime sont possibles suivant qu'il existe, sur le site de recombinaison, quelques nucléotides communs aux deux séquences recombinées (figure 1B), aucun nucléotide commun aux deux séquences réarrangées (figure 1C) ou des nucléotides supplémentaires n'étant auparavant présents sur aucune des deux

séquences impliquées dans la jonction de recombinaison (figure 1D). L'expression « très peu d'homologie » est imprécise et n'est généralement définie que par rapport à l'identité minimale nécessaire à la recombinaison homologue. Dans les cellules de mammifère, par reconstitution d'un gène à partir de deux copies tronquées de ce gène contenant de moins en moins d'identité entre elles, le nombre minimal de nucléotides consécutifs identiques nécessaire à la recombinaison homologue est estimé à quatorze pour la recombinaison intra-

RÉFÉRENCES

1. Meuth M. Illegitimate recombination in mammalian cells. In : Berg, Howe, eds. *Mobile DNA*. Washington DC : American Society of Microbiology, 1989 : 833-60.
2. Roth D, Wilson J. Illegitimate recombination in mammalian cells. In : Kucherlapati R Smith, eds. *Genetic recombination*. Washington DC : American Society of Microbiology, 1988 : 621-53.
3. Ehrlich SD, Bierne H, d'Alencon E, Vilette D, Petranovic M, Noiro P, Michel B. Mechanisms of illegitimate recombination. *Gene* 1993 ; 135 : 161-6.
4. Rubnitz J, Subramani S. The minimum amount of homology required for homologous recombination in mammalian cells. *Mol Cell Biol* 1984 ; 4 : 2253-8.
5. Ayares D, Chekuri L, Kyu-Young S, Kucherlapati R. Sequence homology requirements for intermolecular recombination in mammalian cells. *Proc Natl Acad Sci USA* 1986 ; 83 : 5199-203.
6. Brouillette S, Chartrand P. Intermolecular recombination assay for mammalian cells that produces recombinants carrying both homologous and non-homologous junctions. *Mol Cell Biol* 1987 ; 7 : 2248-55.
7. Nicholls RD, Fischel-Ghodsian N, Higgs DR. Recombination at the human α -globin gene cluster : sequence features and topological constraints. *Cell* 1987 ; 49 : 369-78.
8. Vanin EF, Henthorn PS, Kiuoussis D, Grosveld F, Smithies O. Unexpected relationships between four large deletions in the human β -globin gene cluster. *Cell* 1983 ; 35 : 701-9.
9. Henthorn PS, Mager DL, Huisman THJ, Smithies O. A gene deletion ending within a complex array of repeated sequences 3' to the human β -globin gene cluster. *Proc Natl Acad Sci USA* 1986 ; 83 : 5194-8.
10. Hu X, Ray PN, Worton RG. Mechanisms of tandem duplication in the Duchenne muscular dystrophy gene include both homologous and nonhomologous intrachromosomal recombination. *EMBO J* 1991 ; 10 : 2471-7.
11. Lehrman MA, Goldstein JL, Russel DW, Brown MS. Duplication of seven exons in LDL receptor gene caused by Alu-Alu recombination in a subject with familial hypercholesterolemia. *Cell* 1987 ; 48 : 827-35.
12. Michelin S, Daya-Grosjean L, Sureau F, Said S, Sarasin A, Suarez HG. Characterization of a *C-MET* proto-oncogene activated in human xeroderma pigmentosum cells after treatment with N-methyl-N'-nitro-N-nitrosoguanidine (MNNG). *Oncogene* 1993 ; 8 : 1983-91.
13. Brown L, Cheng JT, Chen Q, Siciliano MJ, Crist W, Buchanan G, Baer R. Site-specific recombination of the *tal-1* gene is a common occurrence in human T cell leukemia. *EMBO J* 1990 ; 9 : 3343-51.

elle représente un compétiteur majeur vis-à-vis des processus de recombinaison homologue nécessaires dans certains types d'expériences, telles que la thérapie génique.

Conséquences biologiques et mécanismes de la recombinaison illégitime

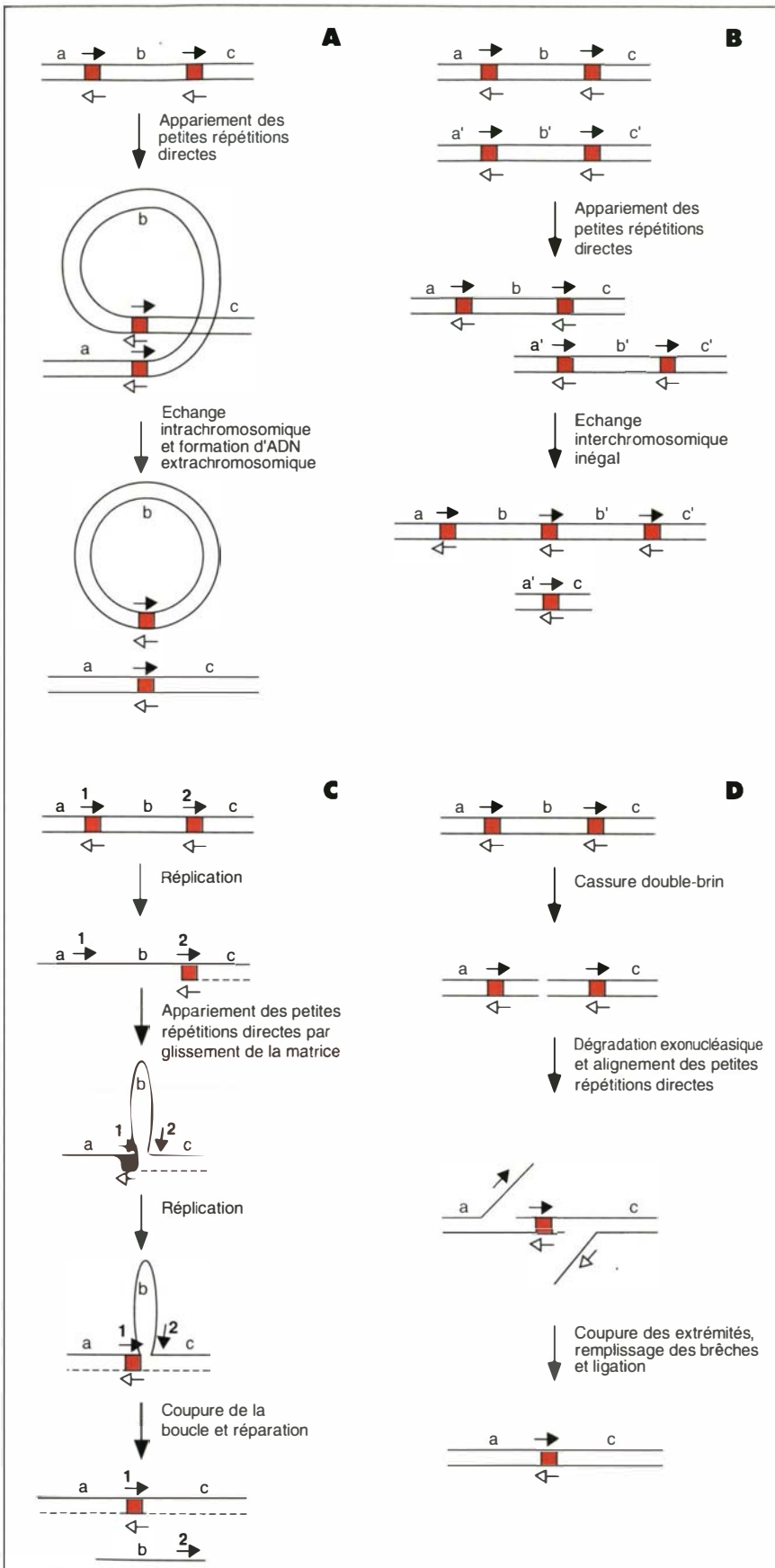
Une des hypothèses quant aux conséquences biologiques de la recombinaison illégitime est qu'elle permettrait une meilleure survie des cellules endommagées grâce à l'élimination des cassures double-brin accidentelles de la chaîne nucléotidique, consécutives à des erreurs du métabolisme normal de l'ADN (réplication, transcription, réparation). Elle peut être considérée comme un mécanisme de sauvegarde cellulaire permettant de réparer les coupures double-brin accidentelles au prix de la production de chromosomes endommagés.

La recombinaison illégitime est impliquée dans différents types de réarrangements (délétions, insertions, translocations, amplifications géniques, intégrations virales). Cela implique que, sous le terme « recombinaison illégitime », sont regroupés des phénomènes ne faisant pas nécessairement appel à des mécanismes moléculaires identiques. Un modèle de recombinaison développé pour un type de remaniement ne peut pas être automatiquement étendu à l'ensemble des phénomènes de recombinaison non homologue. Il n'existe donc pas de modèle général de la recombinaison illégitime.

Un processus en deux étapes a cependant été proposé [2], dans lequel des extrémités d'ADN sont d'abord créées puis réunies dans un second temps. Il pourrait donc exister des points communs aux différents types de réarrangements, comme des facteurs provoquant la cassure de la chaîne nucléotidique ou ceux permettant la réunion de deux extrémités créant une nouvelle séquence d'ADN. Par ailleurs, en s'appuyant sur l'analyse précise des jonctions impliquées dans les délétions, Meuth [1] a constaté que

celles-ci présentaient, le plus souvent, des nucléotides communs aux deux séquences parentales réarrangées (2-7 nucléotides). Il a proposé quatre mécanismes de délétions simples (figure 2) basés sur l'appariement de petites séquences, répétées directes et éloignées, qui provoquerait la perte du matériel génétique localisé entre ces répétitions et la conservation, au site de recombinaison, de l'une de ces répétitions. On trouve ainsi l'échange intrachromosomique conduisant à la formation d'ADN circulaires extrachromosomiques (figure 2A), le *crossing-over* inégal entre courtes répétitions se trouvant sur deux brins d'ADN non liés (figure 2B), l'erreur d'appariement de répétitions par glissement de l'un des brins de la double hélice lors de la réplication (figure 2C) et la dégradation exonucléasique de part et d'autre d'un point de rupture de la double hélice suivie d'une réassociation au niveau de petites répétitions directes (figure 2D).

L'analyse expérimentale de la recombinaison illégitime dans les cellules de mammifère nécessite, d'une part, l'isolement et la caractérisation moléculaire des jonctions de recombinaison engendrées par différents types de réarrangement, dans différents systèmes et, d'autre part, la comparaison des séquences impliquées directement au site de recombinaison et dans les régions proches de ce site afin de déterminer des facteurs communs susceptibles d'être responsables de la recombinaison non homologue. Les séquences du site de recombinaison et celles des deux régions parentales avant le réarrangement doivent donc être déterminées afin de décider si le réarrangement est issu d'une recombinaison illégitime. Du fait de la complexité du génome des mammifères, les modèles d'étude sont largement dépendants du développement de techniques d'analyse moléculaire des gènes eucaryotes et des difficultés à cloner les régions génomiques réarrangées. C'est pourquoi l'étude des réarrangements non homologues est, à l'heure actuelle, limitée à des gènes de séquence nucléotidique connue dont l'altération ou l'activation provoque l'apparition d'un phénotype identifiable et sélectionnable.



Les modèles d'étude des jonctions de recombinaison illégitime

Des jonctions de recombinaison illégitime ont été mises en évidence, dans les cellules germinales, lors de réarrangements génomiques ayant entraîné l'apparition de maladies génétiques héréditaires. Cela a tout d'abord été étudié chez les patients atteints de thalassémie [7, 8] ou de persistance héréditaire de l'hémoglobine fœtale [9]. Plus récemment, la disponibilité de sondes moléculaires pour un nombre plus important de maladies génétiques (hypercholestérolémie familiale, dystrophie musculaire de Duchenne...) a permis d'accroître le nombre de données concernant ces réarrangements qui sont principalement des délétions. Bien que deux cas de recombinaison homologue aient été décrits [10, 11], ce sont essentiellement des réarrangements par recombinaison illégitime qui seraient à l'origine de l'apparition de maladies génétiques, lorsque celles-ci ne sont pas dues à des mutations ponctuelles ou à des expansions de courtes séquences répétées.

Certaines tumeurs peuvent être liées à des translocations d'oncogènes dominants (gènes *c-MYC* et *c-ABL*) ou de gènes potentiellement oncogéniques (gènes *TAL-1* et *MET* [12]). Les gènes des immunoglobulines et des récepteurs des lymphocytes T sont fréquemment impliqués dans ces réarrangements en tant que

Figure 2. Modèles de formation des délétions chromosomiques simples. Ces modèles sont basés sur la présence des petites séquences répétées directes (■) sur les deux partenaires de la recombinaison. Le trait en pointillé indique le brin en cours de réplication. A : échange intrachromosomique et formation d'ADN circulaire extrachromosomique ; B : crossing-over inégal entre courtes répétitions sur deux brins d'ADN non liés ; C : erreur d'appariement de répétitions par glissement de l'un des brins lors de la réplication ; D : dégradation exonucléasique de part et d'autre d'un point de rupture double-brin et réassociation au niveau de petites répétitions directes. (D'après Meuth [1].)

RÉFÉRENCES

14. Hartl P, Lipp M. Generation of a variant t(2; 8) translocation of Burkitt's lymphoma by site-specific recombination via the kappa light-chain joining signals. *Mol Cell Biol* 1987; 7 : 2037-45.
 15. Tsujimoto Y, Gorham J, Cossman J, Jaffe E, Croce CM. The t(14; 18) chromosome translocations involved in B-cell neoplasms result from mistakes in VDJ joining. *Science* 1985; 229 : 1390-3.
 16. Gellert M. Molecular analysis of V(D)J recombination. *Adv Rev Genet* 1992; 22 : 425-46.
 17. Lutzker SG, Alt FW. Immunoglobulin heavy-chain class switching. In : Berg, Howe, eds. *Mobile DNA*. Washington DC : American Society of Microbiology, 1989 : 693-714.
 18. Aurias A, Dutrillaux B. Probable involvement of immunoglobulin superfamily genes in most recurrent chromosomal rearrangement from ataxia-telangiectasia. *Hum Genet* 1986; 72 : 210-4.
 19. Russo G, Isobe M, Pegoraro L, Finan J, Nowell PC, Croce CM. Molecular analysis of a t(7; 14) (q45; q32) chromosome translocation in a T cell leukemia of a patient with ataxia telangiectasia. *Cell* 1988; 53 : 137-44.
 20. Fuscoe JC, Zimmerman LJ, Lippert MJ, Nicklas JA, O'Neill JP, Albertini RJ. V(D)J recombinase-like activity mediates *HPRT* gene deletion in human fetal T-lymphocytes. *Cancer Res* 1991; 51 : 6001-5.
 21. Hyrien O, Debatisse M, Buttin G, Robert de Saint-Vincent B. A hotspot for novel amplification joints in a mosaic of Alu-like repeats and palindromic A+T-rich DNA. *EMBO J* 1987; 6 : 2401-8.
 22. Morris T, Thacker J. Formation of large deletions by illegitimate recombination in the *HPRT* gene of primary human fibroblasts. *Proc Natl Acad Sci USA* 1993; 90 : 1392-6.
 23. Stary A, James MR, Sarasin A. High recombination rate of an Epstein-Barr virus/simian virus 40 hybrid shuttle vector in human cells. *J Virol* 1989; 63 : 3837-43.
 24. Matsuo T, Heller M, Petti L, O'Shiro E, Kieff E. Persistence of the entire Epstein-Barr virus genome integrated into human lymphocyte DNA. *Science* 1984; 226 : 1322-5.
 25. Buendia MA. Hepatitis B viruses and hepatocellular carcinoma. *Adv Cancer Res* 1992; 59 : 167-224.
 26. Skalka AM. Integrative recombination of retroviral DNA. In : Kucherlapati R, Smith, eds. *Genetic recombination*. Washington DC : American Society of Microbiology, 1988 : 701-24.
 27. Samulski RJ, Zhu X, Xiao X, Brook JD, Housman DE, Epstein N, Hunter LA. Targeted integration of adeno-associated virus (AAV) into human chromosome 19. *EMBO J* 1991; 10 : 3941-50.
 28. Derby Shire MK, Epstein LH, Young CSH, Munz PL, Fishel R. Nonhomologous recombination in human cells. *Mol Cell Biol* 1994; 14 : 156-69.
- séquences réceptrices ou activatrices de l'oncogène. L'analyse nucléotidique des séquences impliquées a permis de suggérer [13-15] que ces recombinaisons non homologues seraient dues à des erreurs du mécanisme normal de recombinaison spécialisée des immunoglobulines lors de la différenciation lymphoïde : la recombinaison somatique spécifique du site V-(D)-J [16] et celle intervenant dans les réarrangements des chaînes lourdes des immunoglobulines [17]. Bien qu'il y ait interférence entre recombinaison spécialisée et recombinaison illégitime, il faut signaler que les sites de recombinaison mis en évidence dans les cellules néoplasiques sont différents des sites normaux de recombinaison spécialisée. De plus, les séquences « signal » complètes de la recombinaison spécialisée ne sont pas nécessairement retrouvées sur les deux séquences parentales impliquées dans la recombinaison illégitime. Des erreurs lors du processus normal de recombinaison des immunoglobulines semblent également être la cause de translocations chromosomiques chez les malades atteints d'ataxie télangiectasie (AT) [18, 19]. L'analyse moléculaire d'une translocation chez un malade AT atteint d'une leucémie des cellules T a permis d'identifier des séquences « signal » de la recombinaison spécialisée V-(D)-J dans les deux régions impliquées dans la translocation [19]. Les erreurs de ce système de recombinaison ne sont pas limitées aux gènes des immunoglobulines. En effet, en étudiant les délétions dans le gène *HPRT* humain (hypoxanthine-guanine phosphoribosyltransférase) dans des lymphocytes T fœtaux, Fuscoe *et al.* [20] ont montré que les réarrangements non homologues dans ce gène « de ménage » seraient également provoqués par les recombinaisons impliquées dans le système de recombinaison V-(D)-J. Dans les cellules de mammifère en culture, la sélection de phénotypes particuliers (résistance à des médicaments, transformation) a permis d'identifier, dans des gènes cellulaires [1, 21, 22], des réarrangements génomiques dans lesquels des jonctions de recombinaison illégitime ont été décrites. L'utilisation de sondes exogènes, comme des virus ou des plasmides, a permis de déterminer un grand nombre de jonctions de recombinaison illégitime, soit de façon expérimentale dans les cellules en culture [23], soit au sein d'un organisme comme c'est le cas de l'intégration de certains virus dans les cellules humaines (virus d'Epstein-Barr [24], virus de l'hépatite B [25]). Ainsi, l'insertion de séquences virales dans le génome des mammifères s'effectue généralement de manière non homologue au hasard dans le génome [26]. Il faut, cependant, noter le cas de l'intégration des AAV (*adeno-associated virus*) dans une région spécifique du génome humain au niveau du chromosome 19 dans la bande 19q13.4-ter. L'analyse moléculaire des séquences impliquées a révélé que cette intégration ciblée s'effectuait néanmoins par recombinaison non homologue [27]. Enfin, l'utilisation d'extraits nucléaires de cellules humaines a permis de caractériser précisément les propriétés de refermeture par recombinaison illégitime de molécules contenant des coupures double-brin à des sites connus et d'isoler certaines protéines impliquées dans ce processus [28, 29].

Facteurs communs aux jonctions de recombinaison illégitime

Au point exact des sites de recombinaison, il n'existe pas de séquence nucléotidique impliquée de manière préférentielle aux jonctions de recombinaison non homologue [28-31]. Les cellules de mammifère ont en effet une capacité remarquable à joindre n'importe quel type d'extrémités. Il n'est donc pas étonnant qu'aucune séquence primaire consensus n'ait été rapportée, quel que soit le type de réarrangement ou le modèle d'étude. Au niveau des régions impliquées, certains motifs sont retrouvés aux sites de recombinaison ou dans les régions proches de ces sites, suggérant que toutes les séquences d'ADN ne sont pas équivalentes vis-à-vis de la recombinaison illégitime. Cependant, la présence de ces motifs n'est une condition ni nécessaire, ni suffisante

pour créer un site de recombinaison illégitime puisque leur présence est loin d'être évidente dans tous les cas. Ainsi, comme l'illustre le *Tableau I*, les membres de la famille de séquences répétées de type *Alu* sont fréquemment retrouvés aux sites de recombinaison, soit sur l'une des deux séquences impliquées, soit sur les deux et dans les différents modèles d'étude des réarrangements par

recombinaison illégitime. La famille des séquences courtes répétées de type *Alu* [32] fait partie des séquences dispersées moyennement répétées dans le génome de certains mammifères. Ces répétitions de 300 pb sont retrouvées en moyenne toutes les 5 000 paires de bases dans le génome humain. Elles sont essentiellement distribuées au hasard le long du génome, bien qu'il existe des

concentrations de répétitions *Alu* dans certaines régions. Les répétitions de type *Alu* sont notamment caractérisées par la présence d'une unité promotrice reconnue par l'ARN polymérase III et d'une queue poly(A) à leur terminaison ; elles sont, de plus, bordées par de courtes séquences répétées directes. Ces caractéristiques ont permis de suggérer que ces répétitions seraient des élé-

Tableau I
MOTIFS COMMUNS A CERTAINES SÉQUENCES RÉARRANGÉES
PAR RECOMBINAISON ILLÉGITIME

Motifs	Types de réarrangement	Gène ou sonde	Maladie	Références
Séquences répétées de type <i>Alu</i>	Délétion	α -globine β -globine	α -thalassémie β -thalassémie	[7] [8], [9]
	Insertion	Récepteur des LDL Inhibiteur du composé C1 du complément <i>HPRT</i>	Hypercholestérolémie familiale Angioedema héréditaire	[33], [35] [34]
		EBV Plasmide SV40/EBV Plasmide SV40 Plasmide Ad-5/SV40		[22] [24] [53] [37] [54]
Excision Amplification	Plasmide SV40 <i>ampd</i>		[31] [21]	
Séquences palindromiques	Délétion	α -globine β -globine <i>aprt</i>	α -thalassémie β -thalassémie	[7] [9] [43]
	Insertion Excision Amplification	Polyome Plasmide SV40 <i>ampd</i>		[39] [31] [21]
Séquences alternées de purines-pyrimidines	Délétion Insertion Excision Translocation	SV40 <i>aprt</i> Plasmide SV40 <i>BCL-2</i>	Lymphome folliculaire Leucémie aiguë des cellules T	[44] [43] [31] [40] [41]
	Site de clivage des topo-isomérase I et II	Délétion Insertion	SV40 Gène de la DMD Polyome SV40	Dystrophie musculaire de Duchenne
Translocation Amplification		ALL-1/AF-9 <i>ampd</i>	Leucémie myéloïde aiguë	[56] [21]
Séquences riches en nucléotides AT	Insertion	Plasmide SV40		[37]
	Amplification	<i>ampd</i>		[21]

LDL : lipoprotéines de faible densité ; HPRT : gène de l'hypoxanthine phosphoribosyl transférase.
EBV : virus d'Epstein-Barr ; SV40 : virus simien 40 ; Ad-5 : adénovirus-5.
ampd : gène de l'adénylate désaminase ; aprt : gène de l'adénine phosphoribosyl transférase.

RÉFÉRENCES

29. Thacker J, Chalk J, Ganesh A, North P. A mechanism for deletion formation in DNA by human cell extracts: the involvement of short sequence repeats. *Nucleic Acids Res* 1992; 20: 6183-8.
30. Roth DB, Wilson JH. Nonhomologous recombination in mammalian cells: role for short sequence homologies in the joining reaction. *Mol Cell Biol* 1986; 6: 4295-304.
31. Stary A, Sarasin A. Molecular analysis of DNA junctions produced by illegitimate recombination in human cells. *Nucleic Acids Res* 1992; 20: 4269-74.
32. Deininger PL. Sines: short interspersed repeated DNA elements in higher eucaryotes. In: Berg, Howe, eds. *Mobile DNA*. Washington DC: American Society of Microbiology, 1989: 619-36.
33. Lehrman MA, Russell DW, Goldstein JL, Brown MS. Exon-Alu recombination deletes 5 kilobases from the low density lipoprotein receptor gene, producing a null phenotype in familial hypercholesterolemia. *Proc Natl Acad Sci USA* 1986; 83: 3679-83.
34. Stoppa-Lyonnet D, Carter PE, Meo T, Tosi M. Clusters of intragenic Alu repeats predispose the human C1 inhibitor locus to deleterious rearrangements. *Proc Natl Acad Sci USA* 1990; 87: 1551-5.
35. Benlian P, Loux N. Hétérogénéité des mutations du récepteur LDL dans l'hypercholestérolémie familiale. *médecine/sciences* 1991; 7: 1052-60.
36. Calabretta B, Robberson DL, Barrera-Saldana HA, Lambrou TP, Saunders GF. Genome instability in a region of human DNA enriched in Alu repeat sequences. *Nature* 1982; 296: 219-25.
37. Murnane JP, Yezzi MJ, Young BR. Recombination events during integration of transfected DNA into normal human cells. *Nucleic Acids Res* 1990; 18: 2733-8.
38. Mézard C, Pompon D, Nicolas A. Recombination between similar but not identical DNA sequences during yeast transformation occurs within short stretches of identity. *Cell* 1992; 70: 659-70.
39. Streuli CH, Krauzewicz NS, Griffin BE. Recombination resulting in unusual features in the polyomavirus genome isolated from a murine tumor cell line. *J Virol* 1990; 64: 3570-80.
40. Adachi M, Tsujimoto Y. Potential Z-DNA elements surround the breakpoints of chromosome translocation within the 5' flanking region of *bcl-2* gene. *Oncogene* 1990; 5: 1653-7.
41. Boehm T, Mengle-Gaw L, Kees UR, Spurr N, Lavenir I, Forster A, Rabbitts TH. Alternating purine-pyrimidine tracts may promote chromosomal translocations seen in a variety of human lymphoid tumours. *EMBO J* 1989; 8: 2621-31.
42. Holliday R. Untwisting B-Z DNA. *Trends Genet* 1989; 5: 355-6.
- ments mobiles faisant intervenir un mécanisme de rétrotransposition [32]. En raison du nombre élevé de copies de ces répétitions le long du génome humain, leur implication lors de la recombinaison est controversée. Divers auteurs leur attribuent un rôle dans le processus de recombinaison [7, 33, 34], tandis que d'autres ne veulent pas généraliser leur importance [8]. Cependant, les cartes de distribution des délétions par rapport aux répétitions *Alu* dans le gène de l' α -globine [7] et dans le gène du récepteur des lipoprotéines de faible densité [35], montrent que les points de cassure des remaniements coïncident le plus souvent avec la position des séquences répétées *Alu*. Une instabilité génétique a été observée dans des régions riches en séquences *Alu* [36]. Du fait de l'hypothèse de la mobilité de ces séquences, les régions aptes à intégrer des séquences *Alu* pourraient correspondre elles-mêmes à des zones préférentielles de recombinaison. Dans les régions réarrangées, les séquences *Alu* seraient révélatrices de l'instabilité de la région plutôt qu'instigatrices de la recombinaison [37]. Il faut ajouter que la recombinaison n'implique pas nécessairement un alignement entre les répétitions *Alu* puisque les deux séquences parentales impliquées ont rarement, toutes les deux, ces répétitions au point de recombinaison ou dans les régions proches de ce site. Cependant, on ne peut pas exclure que, dans certains cas, on soit en présence de recombinaison homologue entre séquences homologues, comme dans la levure [38]. Les séquences répétées inversées adjacentes (séquences palindromiques) qui peuvent former des structures secondaires cruciformes (ou en épingle à cheveux) ont été identifiées (*Tableau I*) dans des séquences parentales impliquées dans le réarrangement mais aussi dans certaines séquences réarrangées. Bien qu'il n'y ait pas de preuve directe pour la formation de ces structures cruciformes *in vivo*, leur participation lors de réarrangements par recombinaison illégitime est fortement suggérée [1, 21, 39]. Différentes hypothèses sur le rôle de ces structures ont été avancées, comme celle de sites d'arrêt pour la progression de la fourche de réplication conduisant à des erreurs au cours de la réplication [21, 39] et celle de sites de reconnaissance préférentielle de nucléases qui pourraient introduire des cassures initiant ainsi la recombinaison [1]. De longues alternances de purines-pyrimidines ont également été mises en évidence aux points de cassure de translocations chromosomiques [40, 41] ainsi que dans d'autres modèles d'étude de la recombinaison illégitime (*Tableau I*). Les séquences alternées purines-pyrimidines ont ceci de particulier qu'elles peuvent potentiellement adopter une structure en hélice tournant à gauche, ou ADN de forme Z. L'implication de la structure potentielle en ADN-Z dans la recombinaison illégitime résulterait d'un changement de la structure de la chromatine provoquant l'accès de l'ADN aux enzymes impliquées dans la recombinaison [40-42]. La fragilité des séquences aux jonctions ADN-B / ADN-Z pourrait aussi provoquer une augmentation des cassures de l'ADN dans ces régions [43]. L'intervention possible des topo-isomérases (enzymes capables de modifier l'état de superhélicité de l'ADN et qui catalysent des réactions de « coupure-fermeture » des liaisons phosphodiester dans l'ADN) dans les processus de recombinaison illégitime fut suggérée par les travaux de Bullock *et al.* [44]. Ces auteurs ont observé une colocalisation entre les sites de recombinaison produits *in vivo* lors de l'excision de SV40 du chromosome cellulaire de cellules transformées par ce virus et les sites de clivage *in vitro* d'une topo-isomérase eucaryote de type I sur l'ADN de SV40. Fondée sur les séquences consensus des sites de reconnaissance des topo-isomérases déterminés *in vitro*, la présence de sites de clivage de topo-isomérases de type I et II dans les séquences réarrangées par recombinaison illégitime est évoquée dans différentes études (*Tableau I*). Cependant, la comparaison de sites de recombinaison définis *in vivo* avec des sites de topo-isomérisation *in vitro* sur de l'ADN nu n'a pas grand sens. L'implication des topo-isomérases

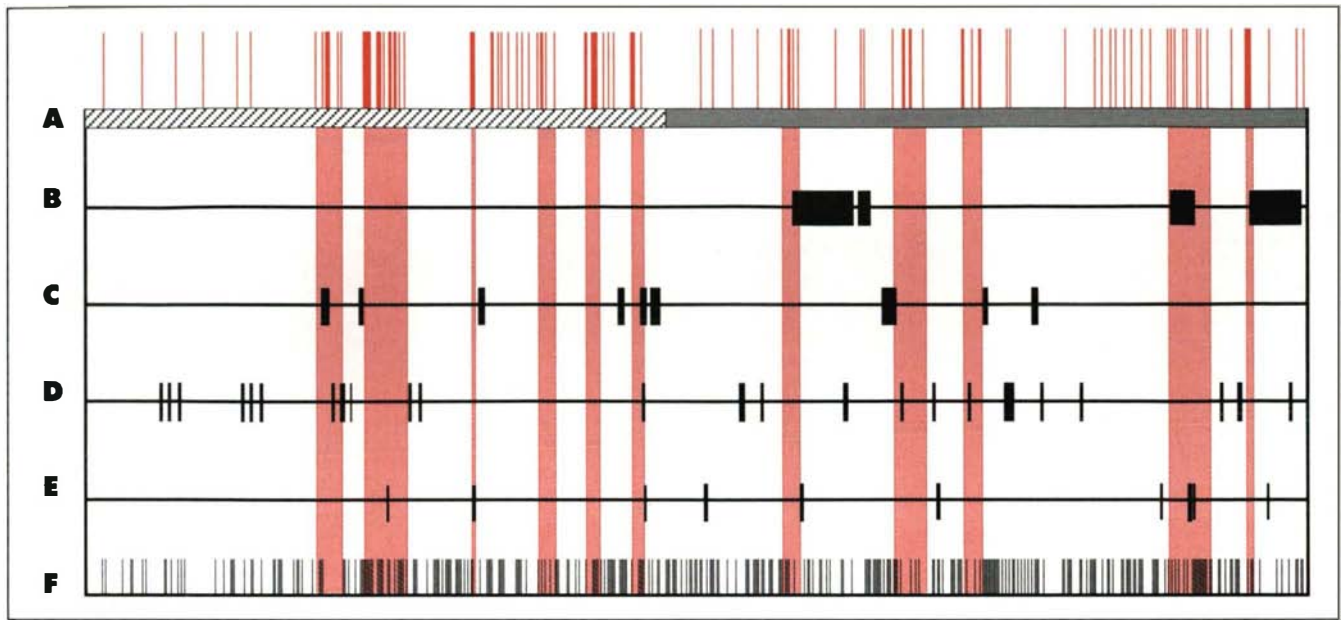


Figure 3. Exemple d'association entre sites de recombinaison illégitime et motifs particuliers d'ADN. Lors de l'excision d'un plasmide SV40 (A, partie hachurée) intégré dans l'ADN génomique humain (A, partie grise), de nombreuses jonctions de recombinaison illégitime ont été mises en évidence. La localisation des séquences impliquées dans ces jonctions est schématisée par les barres rouges en (A). Plusieurs régions présentent un grand nombre de ces séquences recombinées. Ce sont donc des régions de points chauds de recombinaison (bandes rouges verticales). Les motifs retrouvés fréquemment dans des régions de recombinaison illégitime dans les cellules de mammifère ont été recherchés le long des séquences plasmidique et cellulaire. La position des sites de recombinaison est comparée à celle de ces motifs : (B) séquences répétées *Alu* ; (C) séquences palindromiques ; (D) séquences d'homopurine-homopyrimidine ; (E) séquences alternées purine-pyrimidine ; (F) sites de reconnaissance de la topoisomérase I.

dans la recombinaison illégitime est très attrayante puisque ces enzymes pourraient catalyser les deux étapes nécessaires à la recombinaison illégitime (cassure puis réunion) [45]. En comparant 496 jonctions de recombinaison illégitime, Konopka [46] a observé que, dans les 10 paires de bases adjacentes au site de réarrangement, 92 % des jonctions de recombinaison illégitime contenaient les sites préférentiels de reconnaissance *in vitro* pour la topoisomérase I de thymus de veau (CAT, CTY, GTY, RAT où Y est une pyrimidine et R une purine). Il faut, cependant, souligner que la probabilité de trouver des sites de clivage de la topoisomérase I dans une séquence d'ADN est très élevée (un site par sept paires de bases) [1]. La corrélation « recombinaison illégitime-topoisomérase I » reste donc discutable. En analysant la recombinaison *in vitro* entre deux molécules de phage Lambda, Bae *et al.* [47] ont montré que la topoisomérase II de thymus de veau intervenait dans la recombinaison

non homologue. Ils ont proposé un modèle dans lequel l'enzyme lierait l'ADN sous forme d'un complexe dimérique, couperait les deux brins d'ADN, en formant une structure intermédiaire qui pourrait s'associer à un autre complexe dimérique lié à une autre molécule d'ADN. La dissociation de ce complexe tétramérique provoquerait l'échange d'une sous-unité protéique, entraînant un échange entre les deux doubles brins d'ADN. Des segments riches en nucléotides A et T (Tableau 1) ont été identifiés également dans des régions impliquées dans la recombinaison illégitime. En particulier, il a été constaté que les répétitions *Alu*, éléments potentiellement mobiles, sont situés généralement dans des régions de ce type [32] et que l'intégration des rétrovirus dans le génome cellulaire montre également une préférence pour des régions riches en A/T dans la séquence réceptrice [26, 48]. La faible énergie de fusion de l'ADN riche en A/T qui pourrait faciliter

l'ouverture de la matrice double-brin et par conséquent l'entrée des enzymes de recombinaison, est l'hypothèse la plus souvent invoquée sur le rôle de ces régions comme substrat préférentiel de la recombinaison illégitime [21].

Enfin, certaines structures particulières de l'ADN (comme des éléments d'ADN courbé, conformation de l'ADN se produisant dans des séquences contenant des suites d'oligo(dA)-oligo(dT) et provoquant des courbures de l'ADN en déformant l'axe de la double hélice) ont été mises en évidence à proximité de régions génomiques réarrangées par recombinaison illégitime [49].

Conclusion

La caractérisation des sites de recombinaison non homologue et des régions proches de ces sites a permis de mettre en évidence le fait que des séquences particulières, comme les séquences *Alu*, les séquences palindromiques, des alternances purine-

RÉFÉRENCES

43. Nalbantoglu J, Miles C, Meuth M. Insertion of unique and repetitive DNA fragments into the *aprt* locus of hamster cells. *J Mol Biol* 1988 ; 200 : 449-59.
44. Bullock P, Champoux JJ, Botchan M. Association of crossover points with topoisomerase I cleavage sites : a model for nonhomologous recombination. *Science* 1985 ; 230 : 954-8.
45. Duguet M, Riou JF. De la topologie de l'ADN aux médicaments antibiotiques et anticancéreux. *médecine/sciences* 1994 ; 10 : 963-72.
46. Konopka AK. Compilation of DNA strand exchange sites for non homologous recombination in somatic cells. *Nucleic Acids Res* 1988 ; 16 : 1739-58.
47. Bae YS, Kawasaki I, Ikeda H, Liu LF. Illegitimate recombination mediated by calf thymus DNA topoisomerase II *in vitro*. *Proc Natl Acad Sci USA* 1988 ; 85 : 2076-80.
48. Marin M, Etienne-Julian M, Piechaczyk M, Noël D. L'intégration des rétrovirus : faits et croyances. *médecine/sciences* 1994 ; 10 : 318-24.
49. Milot E, Belmaaza A, Wallenburg JC, Gusew N, Bradley WEC, Chartrand P. Chromosomal illegitimate recombination in mammalian cells is associated with intrinsically bent DNA elements. *EMBO J* 1992 ; 11 : 5063-70.
50. Vijaya S, Steffen DL, Robinson HL. Acceptor sites for retroviral integrations map near DNase I-hypersensitive sites in chromatin. *J Virol* 1986 ; 60 : 683-92.
51. Yunis JJ, Soreng AL, Bowe AE. Fragile sites are targets of diverse mutagens and carcinogens. *Oncogene* 1987 ; 1 : 59-69.
52. Sperry AO, Blasquez VC, Garrard WT. Dysfunction of chromosomal loop attachment sites : illegitimate recombination linked to matrix association regions and topoisomerase II. *Proc Natl Acad Sci USA* 1989 ; 86 : 5497-501.
53. Stary A, Sarasin A. Simian virus 40 (SV40) large T antigen-dependent amplification of an Epstein-Barr virus-SV40 hybrid shuttle vector integrated into the human HeLa cell genome. *J Gen Virol* 1992 ; 73 : 1679-85.
54. Romani M, Casciano I, Querzola F, De Ambrosio A, Siniscalco M. Analysis of a viral integration event in a CG-rich region at the 1p36 human chromosomal site. *Gene* 1993 ; 135 : 153-60.
55. Bodley AL, Huang HC, Yu C, Liu LF. Integration of simian virus 40 into cellular DNA occurs at or near topoisomerase II cleavage hot spots induced by VM-26 (teniposide). *Mol Cell Biol* 1993 ; 13 : 6190-200.
56. Negrini M, Felix CA, Martin C, Lange BJ, Nakamura T, Canaani E, Croce CM. Potential topoisomerase II DNA-binding sites at the breakpoints of a t(9 ; 11) chromosome translocation in acute myeloid leukemia. *Cancer Res* 1993 ; 53 : 4489-92.

pyrimidine, les sites de reconnaissance des topo-isomérases ou les régions riches en nucléotides A et T, étaient fréquemment impliquées. L'hypothèse selon laquelle il existerait des séquences plus susceptibles d'être réarrangées par recombinaison illégitime s'est donc développée [1]. Ces « points chauds » seraient des régions fragiles et potentiellement accessibles aux nucléases capables d'introduire des cassures de la chaîne nucléotidique, points potentiels d'initiation d'événements de recombinaison non homologue.

L'hypothèse de la « configuration ouverte de l'ADN » comme cible des réarrangements est le plus souvent avancée. Cette hypothèse est étayée par plusieurs observations, bien que son implication dans la recombinaison illégitime ne soit pas certaine puisque toutes les séquences de réarrangement n'ont pas été déterminées. Ainsi, les rétrovirus s'intègrent préférentiellement dans des régions sensibles à l'action de la DNase I [50]. De même, il existe une corrélation directe entre les sites chromosomiques fragiles [51] et l'intégration de virus à ADN dans des tumeurs humaines ou dans des cellules transformées en culture. La structure de la chromatine peut également jouer un rôle comme cela a été proposé pour la formation des grandes délétions intrachromosomiques [7, 8]. Les régions permettant l'attachement des boucles de chromosomes à la matrice nucléaire, régions MAR (*matrix association region*) riches en nucléotides A/T et contenant des sites de fixation et de clivage de la topo-isomérase II ainsi que, probablement, les sites d'attachement des enzymes de réplication, de transcription et de réparation, pourraient constituer une des classes de séquences cibles de la recombinaison illégitime comme le suggèrent Sperry *et al.* [52].

En conclusion, on peut envisager que des connaissances plus approfondies sur les mécanismes impliqués dans la recombinaison illégitime permettront de mieux comprendre pourquoi ce type de recombinaison prédomine sur la recombinaison homologue dans les cellules de mammifères et ainsi de mieux orienter les techniques pour l'intégration de gènes dans le cas de la thérapie génique ■

Summary

Illegitimate recombination in mammalian cells

Illegitimate recombination, defined operationally as the cellular process leading to the joining of DNA sequences with little or no homology, is often used in mammalian cells. This recombination process is involved in a variety of DNA rearrangements including deletion, insertion, chromosomal translocation, gene amplification and viral integration. These DNA modifications can produce several cellular abnormalities eventually leading to specific genetic diseases or to carcinogenesis. The biological role and the molecular events that underlie this process are still unclear but illegitimate recombination may arise primarily from errors of DNA metabolism, that is, as mistakes in replication, repair or transcription. Several illegitimate recombination junctions from a variety of sources have been analysed at the DNA sequence level. In some cases, this revealed the presence of some features, such as members of interspersed repetitive DNA families, inverted repeats, alternating purine/pyrimidine runs, topoisomerase cleavage sites. This suggests that specific genomic regions could be prone to be implicated in the illegitimate recombination process.

Remerciements

L'auteur remercie Alain Gentil, Michel Lacasa et Alain Sarasin pour leurs critiques et leurs conseils pendant la préparation du manuscrit de cet article.

TIRÉS A PART

A. Stary.