

Instabilité génomique associée à la rétrotransposition du LINE-1 humain

par Nicolas Gilbert, Aurélien J. Doucet & Alain Bucheton

Institut de Génétique Humaine, UPR 1142, 141, rue de la Cardonille, 34396 Montpellier, cedex 5, France.

Correspondance : Nicolas.Gilbert@igh.cnrs.fr

Reçu le 7 octobre 2004

RÉSUMÉ

Le rétrotransposon LINE-1 (L1) représente environ 17 % du génome humain. Du fait de son grand nombre de copies, il est impliqué dans des remaniements génomiques associés à des événements de recombinaison homologue entre sites hétérologues. De plus, même si la vaste majorité des éléments L1 sont inactifs, certains sont encore capables de se mobiliser par rétrotransposition. L1 est donc un agent mutagène par insertion. De plus, des travaux ont aussi montré que les rétrotransposons actifs

étaient impliqués dans la mobilisation d'autres séquences pour produire des rétro-pseudogènes ou amplifier d'autres séquences répétées. Finalement, des études récentes ont montré que l'élément L1 pourrait être associé à de nouveaux réarrangements génomiques produits lors de l'insertion, tels que des délétions génomiques de grande taille. En conclusion, L1 peut être considéré comme un facteur important qui a affecté et modelé le génome humain par l'intermédiaire de plusieurs mécanismes.

SUMMARY Genomic instability associated with human LINE-1 retrotransposition

LINE-1 (L1) retrotransposon accounts for approximately 17 % of the human genome. Because of the great number of identical copies, L1 can be implicated in genomic rearrangements associated with events of homologous recombination between heterologous sites. Moreover, even if the vast majority of the L1 elements are inactive, some are still able to mobilize themselves by retrotransposition. Thus, L1 is regarded as an insertional mutagenic agent. Moreover,

recent works have shown that active retrotransposons were able to mobilize other sequences to generate retro-pseudogenes or to amplify other repeated sequences. Finally, L1 has been associated recently with new genomic rearrangements generated upon insertions such as large genomic deletions. L1 then can be considered as a major factor that has affected and shaped the human genome through several mechanisms.

INTRODUCTION

Tous les génomes eucaryotes possèdent des séquences mobiles, aussi appelées éléments transposables. Ils sont couramment utilisés comme marqueurs de polymorphisme entre les espèces ou au sein d'une même espèce (Salem *et al.*, 2003 ; Shimamura *et al.*, 1997). Ces séquences répétées sont aussi impliquées dans la plasticité des génomes (Kidwell & Lisch, 2002). On définit deux classes d'éléments mobiles en fonction de l'intermédiaire qu'ils utilisent lors de leur mobilisation. Les éléments de classe I utilisent un intermédiaire ARN et sont appelés rétrotransposons tandis que ceux de classe II transposent directement sous forme ADN et sont appelés transposons (Berg & Howe, 1989). Les rétrotransposons sont à leur tour divisés en trois sous-groupes. Les rétrotransposons à LTR (pour "Long Terminal Repeat"), qui ressemblent aux rétrovirus mais qui ne

contiennent pas, en général, le troisième cadre de lecture responsable du caractère infectieux de ces derniers (Temin, 1985). Les rétrotransposons sans LTR peuvent se subdiviser en deux groupes : les rétrotransposons sans LTR autonomes, appelés aussi LINE (Long Interspersed Element), qui possèdent les activités enzymatiques nécessaires à leur mobilité (Weinet *et al.*, 1986), et les rétrotransposons sans LTR non autonomes, appelés SINE (Short Interspersed Element) ou encore rétroposons, qui sont en général de courtes séquences non codantes d'environ 300 bp (Hutchinson *et al.*, 1989 ; Rogers, 1983).

Chez l'Homme, les éléments transposables représentent plus de 45 % de la masse génomique et la majorité sont de la classe I (Lander *et al.*, 2001). L'élément mobile le plus représenté en masse chez l'Homme est le rétrotransposon LINE-1 (L1) qui constitue près de 17 % de notre génome (Smit, 1999). La vaste majorité de ces

séquences est déficiente pour la rétrotransposition à cause de tronctions en 5' de l'élément, de réarrangements internes ou encore de mutations affectant les cadres de lecture (Grimaldi *et al.*, 1984; Lander *et al.*, 2001). Cependant, on estime que le génome humain haploïde contient environ 80 L1 pleine longueur compétents pour la rétrotransposition (RC-L1 pour rétrotransposition compétent L1, Brouha *et al.*, 2003; Sassaman *et al.*, 1997). La présence en grand nombre de copies ainsi que le mécanisme d'amplification de l'élément L1 ont fait de ce dernier un des acteurs principaux de la variabilité de notre génome.

STRUCTURE DU RÉTROTRANSPON L1

Les éléments RC-L1 possèdent une séquence d'environ 6 Kb (Fig. 1). Celle-ci contient une extrémité 5' non traduite (5' UTR) de 910 nucléotides qui présente une activité promotrice interne (Swergold, 1990), deux cadres ouverts de lecture, ORF1 et ORF2, et une courte séquence 3' non traduite (3' UTR). ORF1 code une protéine de 40 kDa (Holmes *et al.*, 1992) possédant des activités de liaison aux acides nucléiques (Hohjoh & Singer, 1996, 1997). De récents travaux *in vitro* proposent qu'en plus d'être impliquée dans la formation d'un complexe Ribo-Nucléo-Protéique (RNP), la protéine de l'ORF1 aurait une activité chaperonne facilitant l'initiation de l'insertion (Martin & Bushman, 2001). ORF2 code une protéine d'environ 150 kDa présentant trois domaines conservés, un domaine endonucléase EN (Fanning & Singer, 1987; Feng *et al.*, 1996), un domaine transcriptase inverse RT (Mathias *et al.*, 1991) et enfin un domaine riche en cystéine (C) dont on ne connaît pas la fonction (Fanning & Singer, 1987). Finalement, les éléments RC-L1 possèdent dans leur région 3' UTR un segment polypurine riche en guanine suivi d'un signal de polyadénylation. La séquence se termine par une extension polyadénine de taille variable (Dombroski *et al.*, 1991). En général, l'élément L1 est flanqué de part et d'autre du site d'insertion par une courte séquence dupliquée (TSD pour Target Site Duplication) d'environ 7 à 20 nucléotides. Toutefois, il n'est pas rare d'observer des événements de rétrotransposition qui induisent de petites délétions de 2 à 50 pb au site d'insertion (Gilbert *et al.*, 2002; Kondo-Iida *et al.*, 1999; Narita *et al.*, 1993; Symer *et al.*, 2002).

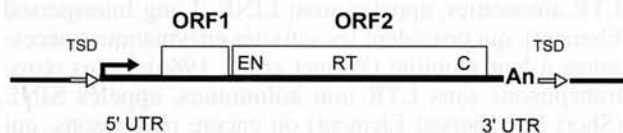


FIG. 1. – Représentation schématique du rétrotransposon L1.

La région 5'UTR contient le promoteur, indiqué par une flèche noire. Les deux cadres de lecture, ORF1 et ORF2, sont en phase et séparés par une courte séquence interstitielle. Les domaines conservés dans ORF2 sont indiqués par EN pour endonucléase, RT pour transcriptase inverse et C pour domaine riche en cystéines. La région 3' UTR est terminée par une queue poly-A de taille variable. La séquence est en général encadrée par la duplication du site d'insertion indiquée par les flèches horizontales blanches (TSD).

MÉCANISME DE RÉTROTRANSPPOSITION DE L1

L'étape initiale de la rétrotransposition est la transcription. Elle s'effectue à partir d'un promoteur interne contenu dans la région 5' UTR de l'élément. Ce promoteur contient vraisemblablement un certain nombre de sites de liaison aux facteurs de transcription dont YY1, SRY et RUNX3 (Swergold, 1990; Tchenio *et al.*, 2000; Yang *et al.*, 2003). Le transcrite est ensuite dirigé vers le cytoplasme où il va être traduit et former un complexe RNP (Hohjoh & Singer, 1996). La formation de ce complexe suit un modèle de *cis*-préférence, c'est-à-dire que les protéines qui sont issues du transcrite de L1 vont se lier de façon préférentielle à l'ARN qui les a produites (Esnault *et al.*, 2000; Wei *et al.*, 2001). Le complexe ainsi formé retourne vers le noyau par un mécanisme encore méconnu. C'est dans le noyau que vont avoir lieu de manière concomitante l'étape de transcription inverse et l'intégration à un nouveau site chromosomique. Ce processus d'insertion appelé TPRT pour "Target-site Primed Reverse Transcription" (Cost *et al.*, 2002) a été originellement décrit pour un autre élément LINE, R2Bm présent dans le génome du ver à soie (Luan *et al.*, 1993; Luan & Eickbush, 1995). Le domaine EN de la protéine issue de l'ORF2 reconnaît un site préférentiel d'insertion non strict, dont le consensus est 5'-TTTTA, et va couper l'ADN simple brin entre le T et le A. Ceci libère une extrémité 3' hydroxyle qui va permettre l'initiation de la transcription inverse au site d'insertion par le domaine RT de la protéine ORF2 et ainsi former le premier brin

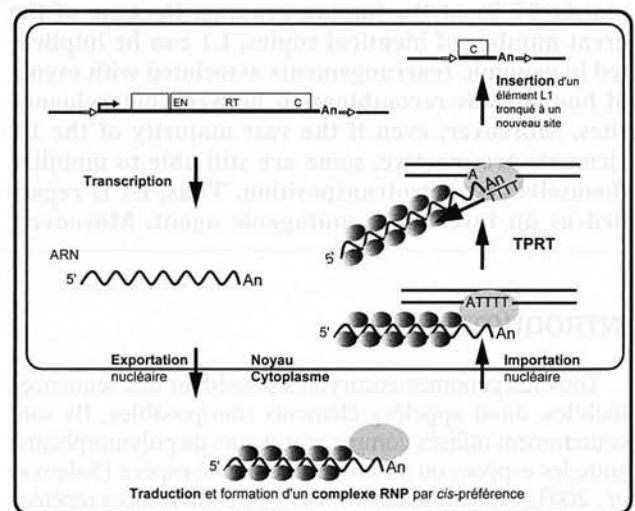


FIG. 2. – Modèle de rétrotransposition de l'élément L1 (d'après Moran & Gilbert, 2002).

Un certain nombre d'étapes restent encore obscures, notamment la composition du complexe RNP et son entrée dans le noyau qui peut être soit active et nécessite alors un signal de localisation, soit passive lors de la dissolution de la membrane nucléaire pendant la division cellulaire. Les petits ovales représentent la protéine de l'ORF1 et le gros ovale celle de l'ORF2. Dans le noyau, l'insertion s'effectue suivant le modèle de TPRT. Un nouvel événement de rétrotransposition conduit en majorité à un élément L1 tronqué encadré par une duplication du site d'intégration.

d'ADN complémentaire (Fig. 2, Moran & Gilbert, 2002). L'hybridation possible entre les thymines du site d'insertion avec la terminaison polyadénine du transcrit va sans doute favoriser l'initiation de la transcription inverse. La coupure du second brin d'ADN chromosomique, la synthèse du second brin d'ADN complémentaire ainsi que la résolution de l'intégration restent des étapes fuligineuses qui font sans doute intervenir des facteurs de la cellule hôte. Dans la majorité des cas, un nouvel événement d'insertion aboutit à la formation d'une nouvelle copie de l'élément L1 tronqué. Ceci est vraisemblablement dû à une dissociation prématurée du complexe de rétrotransposition lors de la transcription inverse. En effet, la taille moyenne d'une séquence L1 dans le génome humain est d'environ 1 Kb.

MOBILISATION PAR RÉTROTRANSCRIPTION DE SÉQUENCES NON-L1

Même si l'élément L1 montre une forte préférence pour mobiliser son propre ARN, il a été démontré de façon expérimentale qu'il était capable de mobiliser d'autres ARN. En effet, L1 serait responsable de l'amplification à plus d'un million de copies du SINE Alu dans notre génome (Dewannieux *et al.*, 2003). Cela fait de L1 le responsable de près de 30 % de la masse du génome humain. De plus, L1 serait responsable de la formation des rétro-pseudogènes (~ 0.5 % du génome) par transcomplémentation d'ARNm de la cellule hôte (Esnault *et al.*, 2000; Wei *et al.*, 2001). Le mécanisme impliqué pour l'amplification de séquences non-L1 serait identique; cependant nous ne connaissons pas les facteurs qui favoriseraient la mobilisation d'un ARN particulier par rapport à d'autres. Par exemple, dans le cas de la séquence Alu, plusieurs hypothèses ont été proposées pour expliquer le succès de son amplification : la terminaison polyadénylée de la séquence Alu (Boeke, 1997), la capacité de l'ARN de l'élément Alu à former une structure secondaire similaire au snRNA 7SL et ainsi de remplacer ce dernier dans le complexe d'arrêt de traduction SRP9/14 (Maraia, 1995; Sinnett *et al.*, 1991), ou encore le saut de brin au début de la transcription inverse entre l'ARN de L1 et celui de Alu (Deininger *et al.*, 2003). Dans tous les cas, la proximité des deux ARN, L1 et Alu, doit être favorisée au cours d'une des étapes de la rétrotransposition. Un autre exemple est celui de la formation de rétro-pseudogènes U6 (le snRNA U6 faisant partie du complexe majeur d'épissage) souvent associés à une séquence L1 générant un élément chimère U6-L1 au site d'insertion. Le modèle pour expliquer la formation de ces chimeres implique un saut de matrice au cours de la transcription inverse entre l'ARN L1 et celui de U6 (Buzdin *et al.*, 2003).

Il existe deux autres mécanismes par lesquels le rétrotransposon L1 peut mobiliser des séquences non-L1. Ce sont les transductions en 5' ou en 3'. Pour le cas de la transduction en 5', un promoteur situé en amont d'un élément L1 actif peut diriger sa transcription. Ce transcrit

conduit à la mobilisation de la séquence 5' flanquante si la transcription inverse de l'ARN au site d'insertion est complète. Les événements de transduction en 3' sont beaucoup plus fréquents (Kazazian, 2004). Ce phénomène est dû au signal de polyadénylation de L1 qui est faible et qui peut être traversé par l'ARN polymérase et ainsi transcrire la séquence flanquante en 3' jusqu'à un autre signal d'arrêt. Lors de la transcription inverse, ce segment sera copié à un nouveau site d'insertion. Ce mécanisme a potentiellement mobilisé des régions codantes ou aurait permis de déplacer dans notre génome des séquences régulatrices (Moran *et al.*, 1999).

L1 ET LES INSTABILITÉS GÉNÉTIQUES

L'ensemble des séquences répétées mobiles a longtemps été considéré comme ADN « poubelle », ou « junk DNA », car ces séquences ne possèdent aucune fonction établie pour la cellule hôte. Cependant, notamment depuis le séquençage systématique de nombreux génomes, nous sommes en train de comprendre le rôle qu'ont pu avoir l'ensemble des éléments mobiles sur l'organisation et l'évolution des génomes eucaryotes. Ces séquences peuvent en effet agir sur notre génome à plusieurs niveaux. Tout d'abord par insertion *de novo* dans des séquences codantes ou non, ensuite en s'associant à des mécanismes de réparation de l'ADN, ou encore par recombinaison illégitime entre copies d'une même famille.

Instabilité par insertions simples

Du fait de sa mobilité, l'élément L1 est considéré comme un agent mutagène. En effet, à l'heure actuelle, il existe quatorze cas répertoriés où L1 est impliqué dans l'apparition de maladies génétiques (Ostertag & Kazazian, 2001) comme par exemple des cas d'hémophilie (insertion dans le gène du facteur VIII) ou de dystrophie musculaire (insertion dans le gène de la dystrophine). De plus, si on accepte que L1 soit responsable de la mobilisation *de novo* de l'élément Alu, alors L1 est aussi impliqué, de façon indirecte cette fois, dans l'apparition sporadique d'une vingtaine d'autres maladies (Ostertag & Kazazian, 2001). À partir de ces données, la fréquence de rétrotransposition chez l'Homme est estimée à environ 1 événement pour 50 individus (Kazazian, 2004).

Les mutations sont la conséquence d'insertions de rétrotransposons dans des régions codantes, régulatrices ou introniques. Par insertion dans un exon, la séquence codante est modifiée et aboutit à la traduction d'une protéine non fonctionnelle (Fig. 3A). L'élément L1 peut aussi s'insérer dans une région qui contrôle la transcription d'un gène (promoteur ou enhancer) et en empêcher ou en altérer son expression (Ostertag & Kazazian, 2001). Finalement, l'insertion peut avoir lieu dans un intron et induire une modification de l'épissage. L'élément L1 ou une séquence mobilisée par L1 nouvellement insérée peut provoquer l'apparition d'un site accepteur d'épissage cryptique ou encore provoquer des sauts

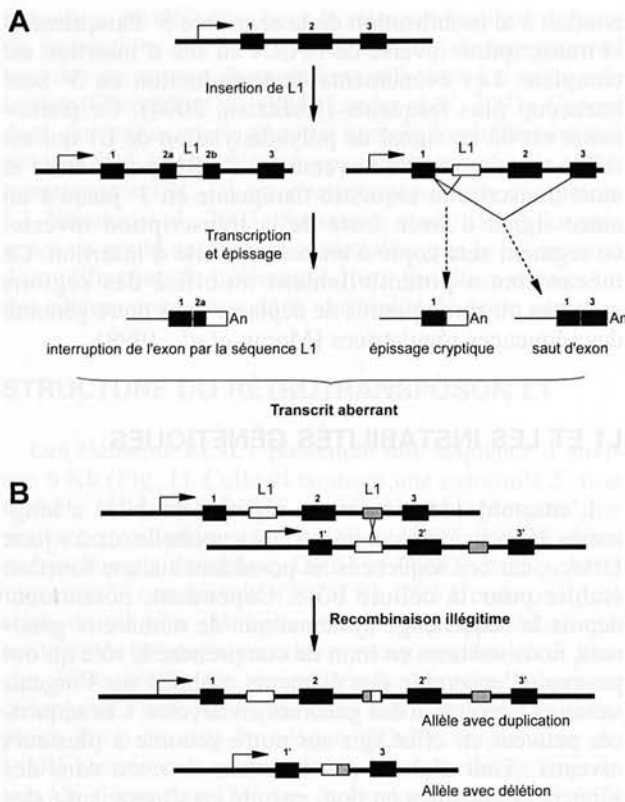


FIG. 3. – Mutations par insertions ou par recombinaisons (d'après Moran & Gilbert, 2002).

A - L'insertion dans un exon conduit à un transcrit aberrant et une protéine non fonctionnelle (partie gauche). L'insertion dans un intron peut provoquer l'apparition d'un épissage cryptique ou le saut d'exon lors de la maturation de l'ARN (partie droite).

B - La recombinaison homologue illégitime se traduit par la formation d'un allèle présentant une duplication et d'un allèle complémentaire présentant une délétion. Si la recombinaison a lieu entre deux éléments se trouvant sur des chromosomes différents, il peut se produire une translocation. Les flèches noires symbolisent les promoteurs. Les exons, notés de 1 à 3, sont représentés par les rectangles noirs et les éléments L1 par des rectangles blancs ou gris.

d'exon lors de l'épissage (Fig. 3A, Batzer & Deininger, 2002 ; Ostertag *et al.*, 2003).

L'insertion d'une séquence répétée à un nouveau locus n'est pas forcément néfaste pour la cellule hôte. En effet, il existe des gènes pour lesquels un élément L1 ou une autre séquence mobilisée par celui-ci, fait partie de la séquence codante (Li *et al.*, 2001). De la même façon, les rétrotransposons peuvent aussi participer à l'activation de la transcription de gènes avoisinants (Ferrigno *et al.*, 2001 ; Ostertag & Kazazian, 2001).

Instabilité par insertions associées à des réarrangements complexes

De récents travaux, effectués grâce à un système d'étude de la rétrotransposition dans des cellules en culture, ont démontré que les insertions d'éléments pouvaient induire des remaniements complexes du génome de la cellule hôte. Dans ces cellules, la rétrotransposition

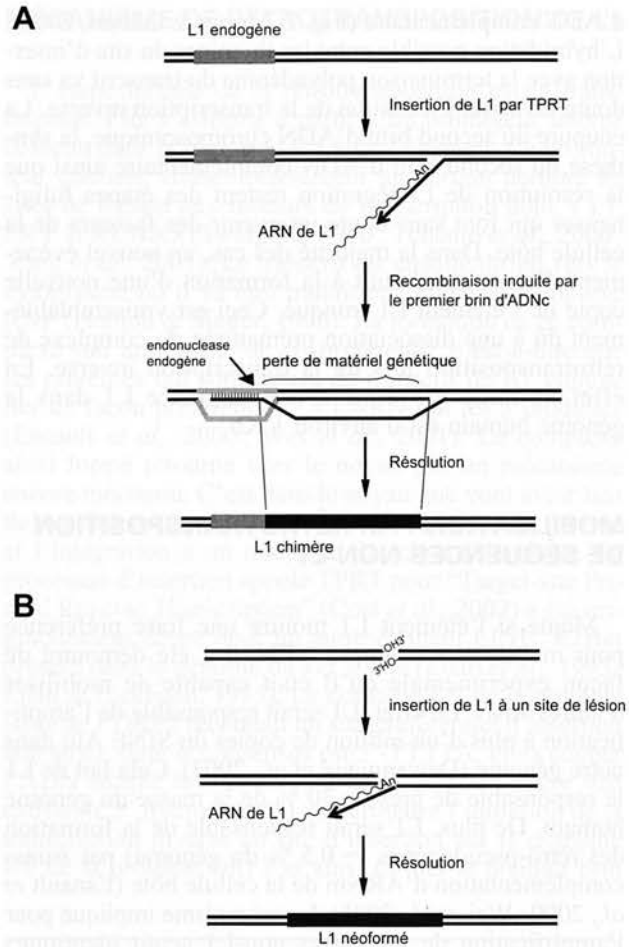


FIG. 4. – Insertion associée à des délétions d'ADN.

A - Délétion d'ADN au site d'insertion générée par homologie de séquence. L'insertion de L1 est initiée par le mécanisme de TPRT. Le premier brin d'ADNc néoformé peut envahir, par homologie de séquence, un élément L1 endogène (rectangle gris) situé en amont du site d'insertion. La résolution de cette structure produit une fusion parfaite entre le L1 endogène et le L1 néoformé (rectangle noir) ainsi qu'une délétion du matériel génétique situé entre le site d'insertion et le L1 endogène. Cette résolution fait intervenir un mécanisme similaire à celui de la réparation de l'ADN par réappariement d'ADN simple-brin complémentaire (SSA).

B - Insertion d'un élément L1 à un site de cassure double brin de l'ADN. Une extrémité 3' hydroxyle de la cassure peut être utilisée pour initier la transcription inverse. Dans des cellules murines en culture déficientes pour le mécanisme de réparation par la réaction d'union d'extrémités non homologues (NHEJ) ce type d'insertion est en général associé à une délétion d'ADN de taille variable au site d'intégration.

de RC-L1 peut créer de larges délétions d'ADN au site d'insertion. Le modèle proposé pour expliquer les pertes de matériel génétique implique des mécanismes de réparation de l'ADN, soit par un système qui nécessite des homologies de séquences, comme par exemple un mécanisme similaire à la réparation d'ADN simple-brin complémentaire (SSA pour "Single-Strand Annealing"), soit par un système de réparation de coupure d'ADN double brin ne nécessitant pas d'homologie de séquence. Dans le premier cas, l'évènement de rétrotransposition pro-

duit, en plus de la délétion, un élément L1 chimère entre la séquence L1 néoformée et un L1 endogène (Fig. 4A, Gilbert *et al.*, 2002; Symer *et al.*, 2002). Dans le second cas, les événements d'insertion semblent être indépendants de l'activité endonucléase de L1. Ces insertions auraient lieu à des sites de cassure double brin de l'ADN et utiliseraient une extrémité 3' hydroxyle libre de la lésion pour initier la transcription inverse. L1 servirait d'outil de réparation de ces coupures (Fig. 4B). Ce phénomène a été observé dans des cellules murines déficientes pour le mécanisme de réparation par la réaction d'union d'extrémités non homologues (NHEJ pour "Non-Homologous End Joining"). Dans ces cellules, les insertions sont fréquemment associées à des pertes d'ADN aux sites d'insertion (Morrish *et al.*, 2002).

Chez la Souris, deux cas de délétions associées à une insertion de L1 ont déjà été observés (Garvey *et al.*, 2002; Kojima *et al.*, 2000). De plus, plusieurs cas similaires mais faisant intervenir l'insertion de séquences Alu ou séquences polyadénylées existent chez l'Homme (Hayakawa *et al.*, 2001; Segal *et al.*, 1999; Su *et al.*, 2000; Wang *et al.*, 2001). Il est probable que la caractérisation de nouvelles maladies génétiques associées à l'insertion d'éléments L1 révèle ces types de réarrangements chez l'Homme.

Instabilité par recombinaison

Les réarrangements que nous venons de décrire sont directement liés au processus de rétrotransposition. Cependant, il existe d'autres réarrangements qui ne sont pas une conséquence directe de ce mécanisme, mais résultent de la nature répétitive de ces éléments. En effet, la recombinaison homologue illégitime entre deux rétrotransposons est favorisée par le grand nombre de copies présentes dans notre génome. Ce type d'évènement a été rarement observé pour l'élément L1 mais est plus fréquent pour le rétrotransposon Alu, sans doute parce que ce dernier se trouve plus fréquemment dans le voisinage de régions codantes. La résolution de la recombinaison produit, sur un allèle, une délétion de matériel génétique et, sur l'allèle complémentaire, une duplication (Fig. 3B). Comme pour les évènements d'insertion, les évènements de recombinaisons peuvent provoquer des maladies génétiques (Batzer & Deininger, 2002; Burwinkel & Kilimann, 1998; Ostertag & Kazazian, 2001; Segal *et al.*, 1999).

CONCLUSION

L'analyse *in silico* des éléments transposables ainsi que le développement récent d'outils moléculaires a permis de mieux évaluer l'impact des séquences mobiles sur les génomes hôtes. Il est bien entendu difficile d'attribuer une fonction précise aux rétrotransposons dans les génomes, mais leur présence semble avoir permis une flexibilité et une variabilité favorisant l'évolution des génomes. Même si son action est potentiellement destructrice à l'échelle d'un individu, le rétrotransposon a

été, et est sans doute encore, bénéfique à l'échelle des espèces pour les raisons décrites ci-dessus. Il existe même un cas particulier où des éléments LINE, *Het-A* et *TART*, sont devenus indispensables à la survie de l'organisme hôte (la *Drosophila*) en remplaçant la fonction de la télomérase (Pardue & DeBaryshe, 2002). Ces éléments se maintiennent donc du fait de l'efficacité de leur propagation mais aussi probablement de leur rôle dans la création de la diversité génétique.

Remerciements. – Nous remercions les membres du laboratoire pour leur évaluation critique du manuscrit. Ce travail a bénéficié d'une aide de l'ARC (Association pour la Recherche sur le Cancer).

BIBLIOGRAPHIE

- Batzer M. A. & Deininger P. L., Alu repeats and human genomic diversity. *Nat. Rev. Genet.*, 2002, 3, 370-379.
- Berg D. E. & Howe M. M., Mobile DNA (Washington, DC, ASM Press), 1989.
- Boeke J. D. LINEs and Alus – the polyA connection. *Nat. Genet.*, 1997, 16, 6-7.
- Brouha B., Schustak J., Badge R. M., Lutz-Prigge S., Farley A. H., Moran J. V. & Kazazian H. H. Jr., Hot L1s account for the bulk of retrotransposition in the human population. *Proc. Natl. Acad. Sci. USA*, 2003, 100, 5280-5285.
- Burwinkel B. & Kilimann M. W., Unequal homologous recombination between LINE-1 elements as a mutational mechanism in human genetic disease. *J. Mol. Biol.*, 1998, 277, 513-517.
- Buzdin A., Gogvadze E., Kovalskaya E., Volchkov P., Ustyugova S., Illarionova A., Fushan A., Vinogradova T. & Sverdlov E., The human genome contains many types of chimeric retrogenes generated through *in vivo* RNA recombination. *Nucleic Acids Res.*, 2003, 31, 4385-4390.
- Cost G. J., Feng Q., Jacquier A. & Boeke J. D., Human L1 element target-primed reverse transcription *in vitro*. *Embo J.*, 2002, 21, 5899-5910.
- Deininger P. L., Moran J. V., Batzer M. A. & Kazazian H. H. Jr., Mobile elements and mammalian genome evolution. *Curr. Opin. Genet. Dev.*, 2003, 13, 651-658.
- Dewannieux M., Esnault C. & Heidmann T., LINE-mediated retrotransposition of marked Alu sequences. *Nat. Genet.*, 2003, 35, 41-48.
- Dombroski B. A., Mathias S. L., Nanthakumar E., Scott A. F. & Kazazian H. H. Jr., Isolation of an active human transposable element. *Science*, 1991, 254, 1805-1808.
- Esnault C., Maestre J. & Heidmann T., Human LINE retrotransposons generate processed pseudogenes. *Nat. Genet.*, 2000, 24, 363-367.
- Fanning T. & Singer M., The LINE-1 DNA sequences in four mammalian orders predict proteins that conserve homologies to retrovirus proteins. *Nucleic Acids Res.*, 1987, 15, 2251-2260.
- Feng Q., Moran J. V., Kazazian H. H. Jr. & Boeke J. D., Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell*, 1996, 87, 905-916.
- Ferrigno O., Virolle T., Djabari Z., Ortonne J. P., White R. J. & Aberdam D., Transposable B2 SINE elements can provide mobile RNA polymerase II promoters. *Nat. Genet.*, 2001, 28, 77-81.
- Garvey S. M., Rajan C., Lemer A. P., Frankel W. N. & Cox G. A., The muscular dystrophy with myositis (*mdm*) mouse mutation disrupts a skeletal muscle-specific domain of titin. *Genomics*, 2002, 79, 146-149.
- Gilbert N., Lutz-Prigge S. & Moran J. V., Genomic deletions crea-



- ted upon LINE-1 retrotransposition. *Cell*, 2002, 110, 315-325.
- Grimaldi G., Skowronski J. & Singer M. F., Defining the beginning and end of KpnI family segments. *Embo J.*, 1984, 3, 1753-1759.
- Hayakawa T., Satta Y., Gagneux P., Varki A. & Takahata N., Alu-mediated inactivation of the human CMP-N-acetylneuraminic acid hydroxylase gene. *Proc. Natl. Acad. Sci. USA*, 2001, 98, 11399-11404.
- Hohjoh H. & Singer M. F., Cytoplasmic ribonucleoprotein complexes containing human LINE-1 protein and RNA. *Embo J.*, 1996, 15, 630-639.
- Hohjoh H. & Singer M. F., Sequence-specific single-strand RNA binding protein encoded by the human LINE-1 retrotransposon. *Embo J.*, 1997, 16, 6034-6043.
- Holmes S. E., Singer M. F. & Swergold G. D., Studies on p40, the leucine zipper motif-containing protein encoded by the first open reading frame of an active human LINE-1 transposable element. *J. Biol. Chem.*, 1992, 267, 19765-19768.
- Hutchison C. A., Hardies S. C., Loeb D. D., Shehee W. R. & Edgell M. H., LINEs and related retrotransposons: long interspersed repeated sequences in the eucaryotic genome. In *Mobile DNA*, Berg D. E. and Howe M. M., eds. (Washington, DC, ASM Press), 1989, pp. 593-617.
- Kazazian H. H. Jr., Mobile elements: drivers of genome evolution. *Science*, 2004, 303, 1626-1632.
- Kidwell M. G. & Lisch D. R., Transposable elements as sources of genomic variation. In *Mobile DNA II*, Craig N. L., Craigie R., Gellert M. & Lambowitz A. M., eds. (Washington, DC, ASM Press), 2002, pp. 59-90.
- Kojima T., Nakajima K. & Mikoshiba K., The disabled 1 gene is disrupted by a replacement with L1 fragment in yotari mice. *Brain Res. Mol. Brain Res.*, 2000, 75, 121-127.
- Kondo-Iida, E., Kobayashi K., Watanabe M., Sasaki J., Kumagai T., Koide H., Saito K., Osawa M., Nakamura Y. & Toda T., Novel mutations and genotype-phenotype relationships in 107 families with Fukuyama-type congenital muscular dystrophy (FCMD). *Hum. Mol. Genet.*, 1999, 8, 2303-2309.
- Lander E. S., Linton L. M., Birren B., Nusbaum C., Zody M. C., Baldwin J., Devon K., Dewar K., Doyle M., FitzHugh W., Funke R., Gage D. et al., Initial sequencing and analysis of the human genome. *Nature*, 2001, 409, 860-921.
- Li W. H., Gu Z., Wang H. & Nekrutenko A., Evolutionary analyses of the human genome. *Nature*, 2001, 409, 847-849.
- Luan D. D. & Eickbush T. H., RNA template requirements for target DNA-primed reverse transcription by the R2 retrotransposable element. *Mol. Cell. Biol.*, 1995, 15, 3882-3891.
- Luan D. D., Korman M. H., Jakubczak J. L. & Eickbush T. H., Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. *Cell*, 1993, 72, 595-605.
- Maraia R. J., The impact of short interspersed elements (SINEs) on the host genome (Austin, Landes Company), 1995.
- Martin S. L. & Bushman F. D., Nucleic acid chaperone activity of the ORF1 protein from the mouse LINE-1 retrotransposon. *Mol. Cell. Biol.*, 2001, 21, 467-475.
- Mathias S. L., Scott A. F., Kazazian H. H. Jr., Boeke J. D. & Gabriel A., Reverse transcriptase encoded by a human transposable element. *Science*, 1991, 254, 1808-1810.
- Moran J. V., DeBerardinis R. J. & Kazazian H. H. Jr., Exon shuffling by L1 retrotransposition. *Science*, 1999, 283, 1530-1534.
- Moran J. V. & Gilbert N., Mammalian LINE-1 retrotransposons and related elements. *Mobile DNA II*, Craig N. L., Craggie R., Geller M. & Lambowitz A., eds. (Washington, DC, ASM Press), 2002, pp. 836-869.
- Morrish T. A., Gilbert N., Myers J. S., Vincent B. J., Stamato T. D., Taccioli G. E., Batzer M. A. & Moran J. V., DNA repair mediated by endonuclease-independent LINE-1 retrotransposition. *Nat. Genet.*, 2002, 31, 159-165.
- Narita N., Nishio H., Kitoh Y., Ishikawa Y., Minami R., Nakamura H. & Matsuo M., Insertion of a 5' truncated L1 element into the 3' end of exon 44 of the dystrophin gene resulted in skipping of the exon during splicing in a case of Duchenne muscular dystrophy. *J. Clin. Invest.*, 1993, 91, 1862-1867.
- Ostertag E. M., Goodier J. L., Zhang Y. & Kazazian H. H. Jr., SVA elements are nonautonomous retrotransposons that cause disease in humans. *Am. J. Hum. Genet.*, 2003, 73, 1444-1451.
- Ostertag E. M. & Kazazian H. H. Jr., Biology of mammalian L1 retrotransposons. *Annu. Rev. Genet.*, 2001, 35, 501-538.
- Pardue M. L. & DeBaryshe P. G., Telomeres and transposable elements. In *Mobile DNA II*, Craig N. L., Craggie R., Gellert M. & Lambowitz A., eds. (Washington, DC, ASM Press), 2002, pp. 870-887.
- Rogers J., Retrotransposons defined. *Nature*, 1983, 301, 460.
- Salem A. H., Ray D. A., Xing J., Callinan P. A., Myers J. S., Hedges D. J., Garber R. K., Witherspoon D. J., Jorde L. B. & Batzer M. A., Alu elements and hominid phylogenetics. *Proc. Natl. Acad. Sci. USA*, 2003, 100, 12787-12791.
- Sassaman D. M., Dombroski B. A., Moran J. V., Kimberland M. L., Naas T. P., DeBerardinis R. J., Gabriel A., Swergold G. D. & Kazazian H. H. Jr., Many human L1 elements are capable of retrotransposition. *Nat. Genet.*, 1997, 16, 37-43.
- Segal Y., Peissel B., Renieri A., de Marchi M., Ballabio A., Pel Y. & Zhou J., LINE-1 elements at the sites of molecular rearrangements in Alport syndrome-diffuse leiomyomatosis. *Am. J. Hum. Genet.*, 1999, 64, 62-69.
- Shimamura M., Yasue H., Ohshima K., Abe H., Kato H., Kishiro T., Goto M., Munechika I. & Okada N., Molecular evidence from retrotransposons that whales form a clade within even-toed ungulates. *Nature*, 1997, 388, 666-670.
- Sinnett D., Richer C., Deragon J. M. & Labuda D., Alu RNA secondary structure consists of two independent 7 SL RNA-like folding units. *J. Biol. Chem.*, 1991, 266, 8675-8678.
- Smit A. F., Interspersed repeats and other mementos of transposable elements in mammalian genomes. *Curr. Opin. Genet. Dev.*, 1999, 9, 657-663.
- Su L. K., Steinbach G., Sawyer J. C., Hindi M., Ward P. A. & Lynch P. M., Genomic rearrangements of the APC tumor-suppressor gene in familial adenomatous polyposis. *Hum. Genet.*, 2000, 106, 101-107.
- Swergold G. D., Identification, characterization, and cell specificity of a human LINE-1 promoter. *Mol. Cell. Biol.*, 1990, 10, 6718-6729.
- Symer D. E., Connelly C., Szak S. T., Caputo E. M., Cost G. J., Parmigiani G. & Boeke J. D., Human L1 retrotransposition is associated with genetic instability *in vivo*. *Cell*, 2002, 110, 327-338.
- Tchenio T., Casella J. F. & Heidmann T., Members of the SRY family regulate the human LINE retrotransposons. *Nucleic Acids Res.*, 2000, 28, 411-415.
- Temin H. M., Reverse transcription in the eukaryotic genome: retroviruses, pararetroviruses, retrotransposons, and retrotranscripts. *Mol. Biol. Evol.*, 1985, 2, 455-468.
- Wang T., Lerer I., Gueta Z., Sagi M., Kadouri L., Peretz T. & Abeliovich D., A deletion/insertion mutation in the BRCA2 gene in a breast cancer family: a possible role of the Alu-polyA tail in the evolution of the deletion. *Genes Chromosomes Cancer*, 2001, 31, 91-95.
- Wei W., Gilbert N., Ooi S. L., Lawler J. F., Ostertag E. M., Kazazian H. H., Boeke J. D. & Moran J. V., Human L1 retrotransposition: cis preference versus trans complementation. *Mol. Cell. Biol.*, 2001, 21, 1429-1439.
- Weiner A. M., Deininger P. & Efstratiadis A., Nonviral retrotransposons: genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information. *Annual Review of Biochemistry*, 1986, 55, 631-661.
- Yang N., Zhang L., Zhang Y. & Kazazian H. H. Jr., An important role for RUNX3 in human L1 transcription and retrotransposition. *Nucleic Acids Res.*, 2003, 31, 4929-4940.