

## Les éléments transposables SINE B2 : source de promoteurs pol II mobiles ? Implications dans l'évolution du génome

Les génomes eucaryotes sont des entités dynamiques très complexes. Seule une fraction de ces génomes consiste en exons codant pour des protéines, alors que la majorité des séquences non codantes sont des éléments répétés. En effet, le génome des mammifères ne contient environ que 2 % d'exons ; 50 % sont constitués d'éléments répétés, les 48 % restant correspondent à des séquences d'ADN unique, dont la plupart ont probablement comme origine des éléments mobiles qui, ayant divergé au cours du temps, ne sont plus reconnaissables [1].

L'abondance des séquences répétées n'a pas d'explication rationnelle évidente et directe d'autant que certains organismes possèdent un génome compact. Ces éléments ont souvent été assimilés à des parasites ou de simples « déchets génomiques » (*junk* ou *selfish DNA*). Cependant, au fur et à mesure des progrès dans le décryptage du génome humain, il est apparu que ces éléments répétés ont contribué largement au dynamisme évolutif du génome qui les héberge [2]. Les éléments transposables constituent la classe majoritaire d'éléments répétés (plus de 90 % chez l'homme), avec principalement les LINE (*long interspaced elements*), qui codent pour les fonctions nécessaires à leur transposition, et les SINE (*short interspaced elements*) qui utilisent la machinerie des LINE pour leur transposition (*m/s* 2001, n° 1, p. 103).

La participation des éléments répétés à la dynamique des génomes est illustrée par le rôle important qu'ils jouent dans la recombinaison homologue inégale, en raison de leurs similitudes de séquences. Ils permettent en effet l'appariement et l'échange de fragments de chroma-

tine non apparentés, causant des délétions ou des duplications de fragments génomiques.

Un autre exemple est le brassage des exons (*genome shuffling*). Les éléments L1 (*human LINE-1*) peuvent incorporer, au cours de leur rétrotransposition, des séquences génomiques situées en aval qui seront aussi insérées à l'endroit même de la nouvelle insertion de ces éléments [3]. On estime que plus de 1 % du génome humain peut être ainsi brassé par la transduction des éléments L1 [4]. L'aspect remarquable de ce mécanisme de « rétro-brassage » d'exons est qu'il ne nécessite pas de séquences homologues, les positions des deux gènes « donneur » et « accepteur » ne sont pas importantes et le gène donneur demeure inaltéré [4]. Enfin, les éléments transposables peuvent aussi moduler la structure et l'expression de gènes « natifs ». Les exemples les plus connus concernent des éléments de contrôle de la transcription et les signaux de polyadénylation [5]. Certains éléments transposables peuvent aussi faire partie intégrante de la séquence codante d'un gène, accroissant encore la variabilité de la structure protéique [5].

Ce qui était jusqu'alors totalement inédit est la capacité potentielle de certains éléments SINE à créer de nouvelles unités de transcription dans les génomes de mammifères [6]. Les SINE sont des composants très abondants de ces génomes, les plus connus étant les éléments Alu chez l'homme. Ils sont transcrits par l'ARN polymérase (pol) III pour donner naissance à de petits transcrits qui ne sont pas traduits. Bien que leur rôle reste inconnu et à démontrer, il est admis que ces éléments ont eu un impact majeur sur la structure et l'évo-

lution du génome. Les familles de SINE consistent généralement en plus de 100 000 membres de longueur identique qui possèdent de 70 % à plus de 98 % d'homologie de séquences. Elles sont apparues indépendamment, assez récemment au cours de l'évolution, chacune à partir d'un gène différent transcrit par l'ARN polymérase III. Par exemple, la famille de séquences Alu chez l'homme dérive de l'ARN 7SL [7] alors que les répétitions B2 forment une famille de SINE plus spécifiques du génome des rongeurs, mais présents aussi chez l'homme, et qui sont dérivés des ARNt (*figure 1*) [8-10].

La transcription de ces éléments par la pol III est augmentée en réponse à un choc thermique et est particulièrement réglée pendant le développement, suggérant un rôle de ces ARN dans ces deux situations [11, 12]. Le niveau de transcription des éléments B2 varie selon les tissus [10]. La transcription dans les testicules et l'augmentation de transcription observée dans les jeunes embryons sont consistants avec la faculté des B2 de passer par la lignée germinale jusqu'à la génération suivante [13]. De plus, il existe une augmentation importante de la transcription par la pol III de ces éléments dans les cellules transformées [11, 12].

Une question non résolue était de savoir pourquoi les éléments B2 ont eu autant de succès dans l'évolution, c'est-à-dire pourquoi ils se sont amplifiés plus activement que les autres familles de SINE. Pour certains, la transcription des éléments B2 semble plus active que celle des autres éléments répétés. Il apparaît en effet que leur promoteur pol III est, soit plus fort, soit moins dépendant de ses séquences flanquantes.

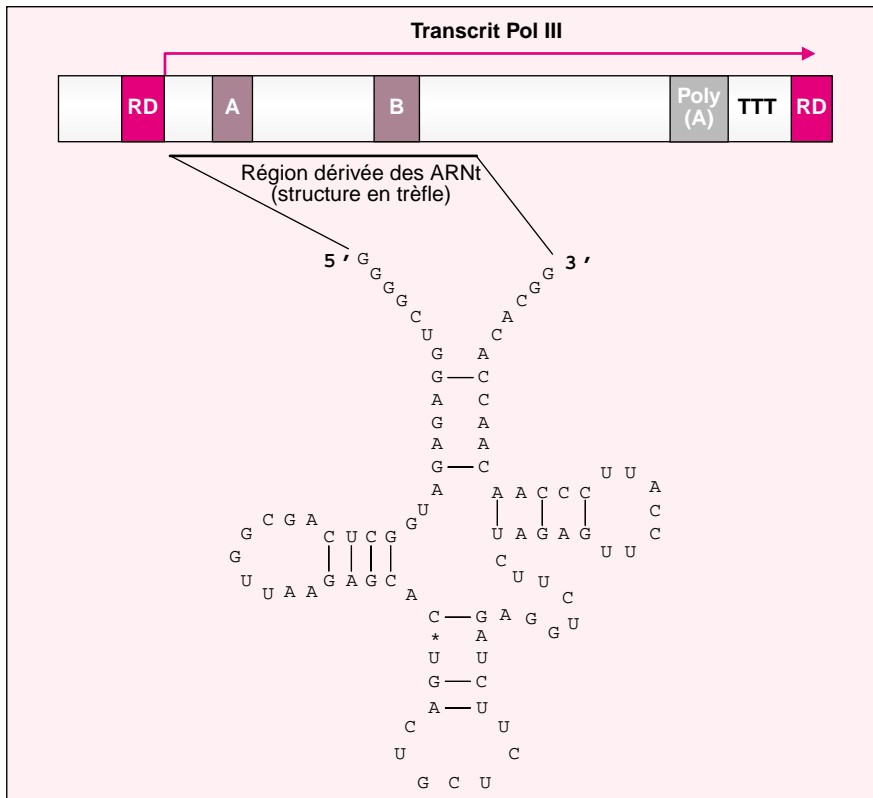


Figure 1. **Structure des éléments B2.** La séquence SINE «générique» contient un promoteur ARN polymérase III (pol III) interne, une séquence riche en A à l'extrémité 3' (sur le brin correspondant au transcrit) et des répétitions directes flanquantes (RD). Ces dernières ne font pas partie des membres de la famille répétée elle-même, mais dérivent de la duplication de la séquence cible au site d'intégration [8]. Les boîtes A et B font partie du promoteur pol III dont la transcription débute à la première base du SINE et continue jusqu'à la rencontre de 4 résidus dT. La plupart des éléments B2 possèdent ce signal poly(A) caractéristique qui, dans certains cas, peut être utilisé lorsqu'il est présent dans la partie 3' non codante d'un gène [9].

De plus, les éléments B2 ont une extrémité 3' d'origine inconnue qui inclut un signal de polyadénylation, l'un ou l'autre de ces éléments pourrait jouer un rôle en augmentant l'efficacité d'une des étapes de la transposition [5]. Une autre possibilité était l'existence d'une sélection positive due à une fonction bénéfique de l'insertion de ces éléments dans l'histoire évolutive.

Au cours de l'analyse de la structure du gène lama3, codant pour la chaîne  $\alpha 3$  des laminines, nous avons identifié un nouveau transcrit ( $\alpha 3C$ ) dont le site d'initiation de la transcription se situe au sein de la région unique d'un élément B2 [6]. De manière remarquable, le transcrit  $\alpha 3C$  est produit par le brin opposé

(antisens) au transcrit pol III dérivé de cet élément. La transcription d' $\alpha 3C$  par la pol II est conduite par un fragment génomique de 70 pb qui appartient exclusivement à la séquence B2 et correspond à la séquence d'origine inconnue des éléments B2. La transcription est déclenchée grâce à une boîte TATA située en amont d'un site d'initiation de la transcription (figure 2). Ce dernier est reconnu par le facteur de transcription USF (*upstream stimulatory factor*) dont l'expression ectopique peut stimuler *in vivo* la transcription par la pol II d' $\alpha 3C$  à partir de l'élément B2. Ces résultats mettent en évidence la présence d'un promoteur pol II dans un élément SINE. Ce promoteur de la pol II n'empêche pas la transcription de l'autre brin par la pol III, étape nécessaire à la transposition de l'élément B2.

La majorité des éléments B2 possède un tel promoteur pol II, et il est très probable que le caractère fonctionnel de ces promoteurs dépende de différents facteurs, dont le contexte chromatinien [6]. La recherche dans les banques de données informatiques nous a permis de trouver de grandes similitudes de séquences entre ces promoteurs pol II des éléments B2 et ceux de gènes connus dont les sites d'initiation de la transcription ont été parfaitement caractérisés. Ceci suggère une origine commune à ces promoteurs et évoque une contribution des éléments de la famille B2 à l'évolution des mammi-

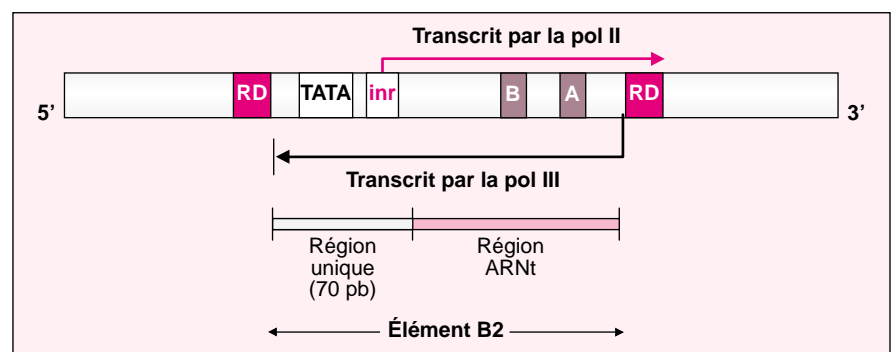


Figure 2. **Identification d'un promoteur pol II au sein d'un élément B2 SINE.** Par convention, le transcrit pol II est représenté de gauche à droite, ce qui place l'élément B2 dans une orientation inverse de celle proposée dans la figure 1, la transcription par la pol II et la pol III se faisant en sens opposés. La région B2 homologue aux ARNt est indiquée, ainsi que le site d'initiation de la transcription (inr) pol II (+1) et la boîte TATA.

fères: la distribution de ces promoteurs mobiles à différents endroits du génome pourrait avoir contribué à la création de nouveaux transcrits, fournissant ainsi le matériel de base pour la sélection naturelle.

Ces résultats fournissent un nouvel argument dans le débat visant à déterminer si les éléments transposables peuvent être des composants bénéfiques aux génomes eucaryotes. Dans l'hypothèse de l'«ADN égoïste», on ne voit pas de façon évidente quel pourrait être le bénéfice, pour un élément B2, de posséder un promoteur pol II dont la transcription s'effectue en sens opposé à sa propre transcription par la pol III. Il faut cependant considérer que la présence de ce nouveau promoteur peut augmenter l'accès à la chromatine des complexes de transcription par la pol III, et inversement. De plus, la dissémination de promoteurs pol II pourrait apporter un avantage pour l'hôte et, par conséquent, augmenter les chances de survie des B2. Pour une raison encore inconnue, l'insertion des SINE se fait dans les régions chromosomiques contenant les gènes actifs et le site d'insertion est riche en A, ce qui pourrait avoir encore augmenté l'impact de leurs insertions.

Enfin, il a souvent été proposé que l'évolution des organismes complexes nécessite moins de protéines nouvelles que de nouveaux schémas de contrôle, principalement par des mécanismes capables de co-régler un ensemble de gènes qui étaient

contrôlés séparément. Il existe en effet une spécificité de réponse à différents éléments régulateurs selon la composition même des éléments de base des promoteurs pol II. La dispersion dans le génome de promoteurs minimums et fonctionnels, ayant la même constitution de base pourrait, ainsi aider la création d'un tel schéma de contrôle.

En conclusions, les éléments transposables, dont on connaît maintenant l'abondance, ont eu une grande influence dans le façonnage du génome: les ARN nucléaires des complexes d'épissage, les rétrovirus, et même la télomérase peuvent tous un jour avoir été des rétrotransposons [14]. Ces dernières années, d'importants progrès dans notre connaissance de la dynamique du génome ont été réalisés, et il apparaît que différents éléments mobiles jouent ainsi un rôle prépondérant dans le processus. En cette période d'après-génome, de plus en plus de biologistes voient les éléments répétés comme un «trésor génomique».

1. Makalowski W. SINEs as a genomic scrap yard: an essay genomic evolution. In: Maraia RJ, ed. *The impact of short interspersed elements (SINEs) on the host genome*. Austin TX: R.G. Landes, 1995: 81-104.
2. Heilig R, Brùls T. Premiers regards sur la séquence du génome humain. *Med Sci* 2001; 17: 299-308.
3. Holmes SE, Dombroski BA, Krebs CM, Boehm CD, Kazazian HH Jr. A new retrotransposable human L1 element from the LRE2 locus on chromosome 1q produces a chimaeric insertion. *Nat Genet* 1994; 7: 143-8.

4. Moran JV, DeBerardinis RJ, Kazazian HH Jr. Exon shuffling by L1 retrotransposition. *Science* 1999; 283: 1530-4.
5. Brosius J. RNAs from all categories generate retrosequences that may be exapted as novel genes or regulatory elements. *Gene* 1999; 238: 115-34.
6. Ferrigno O, Virolle T, Djabari Z, Ortonne JP, White RJ, Aberdam D. Transposable B2 SINE elements can provide mobile RNA polymerase II promoters. *Nat Genet* 2001; 28: 77-81.
7. Schmid C, Maraia R. Transcriptional regulation and transpositional selection of active SINE sequences. *Curr Opin Genet Dev* 1992; 2: 874-82.
8. Daniels GR, Deininger PL. Repeat sequence families derived from mammalian tRNA genes. *Nature* 1985; 317: 819-22.
9. Ullu E, Murphy S, Melli M. Human 7SL RNA consists of a 140 nucleotide middle-repetitive sequence inserted in an alu sequence. *Cell* 1982; 29: 195-202.
10. Ryskov AP, Ivanov PL, Tokarskaya ON, Kramarov DA, Grigoryan MS, Georgiev GP. Major transcripts containing B1 and B2 repetitive sequences in cytoplasmic poly(A)+RNA from mouse tissues. *FEBS Lett* 1985; 182: 73-6.
11. Fornace AJ Jr, Mitchell JB. Induction of B2 RNA polymerase III transcription by heat shock: enrichment for heat shock induced sequences in rodent cells by hybridization subtraction. *Nucleic Acids Res* 1986; 14: 5793-811.
12. Bachvarova R. Small B2 RNAs in mouse oocytes, embryos, and somatic tissues. *Dev Biol* 1988; 130: 513-23.
13. Vasseur M, Condamine H, Duprey P. RNAs containing B2 repeated sequences are transcribed in the early stages of mouse embryogenesis. *EMBO J* 1985; 4: 1749-53.

**Olivier Ferrigno**  
**Daniel Aberdam**

*Inserm U. 385, Faculté de médecine, avenue Valombrose, 06107 Nice Cedex 2, France.*

