

Article

« Ni gènes, ni *junk*, mais des TAR/TUF ! / Neither genes, nor junk, but... TAR/TUF ! »

Jean-Claude Kaplan

M/S : médecine sciences, vol. 21, n° 11, 2005, p. 1005.

Pour citer cet article, utiliser l'information suivante :

URI: <http://id.erudit.org/iderudit/011970ar>

DOI: 10.7202/011970ar

Note : les règles d'écriture des références bibliographiques peuvent varier selon les différents domaines du savoir.

Ce document est protégé par la loi sur le droit d'auteur. L'utilisation des services d'Érudit (y compris la reproduction) est assujettie à sa politique d'utilisation que vous pouvez consulter à l'URI <https://apropos.erudit.org/fr/usagers/politique-dutilisation/>

Érudit est un consortium interuniversitaire sans but lucratif composé de l'Université de Montréal, l'Université Laval et l'Université du Québec à Montréal. Il a pour mission la promotion et la valorisation de la recherche. Érudit offre des services d'édition numérique de documents scientifiques depuis 1998.

Pour communiquer avec les responsables d'Érudit : info@erudit.org

Ni gènes, ni junk, mais des TAR/TUF !

Jean-Claude Kaplan

> L'existence d'une catégorie nouvelle d'ARN ne codant pas de protéines, globalement appelés ARNnc (*noncoding RNA*), est suspectée depuis la découverte de l'ARN interférence, puis des ARNmi. Par ailleurs, l'abondance croissante dans les banques de données des EST (*expressed sequence tags*) et des séquences obtenues par la méthode SAGE, contrastant avec le décompte révisé à la baisse des gènes de protéines [1], laissait supposer que le transcriptome est plus qu'une simple collection d'ARNm codant des protéines. Un autre argument indirect allant dans le même sens est l'abondance des TFBS (*transcription factor binding sites*), au nombre de plusieurs centaines de milliers, en particulier dans les régions dites désertiques (ce qui suggère qu'il y a de la transcription en dehors des gènes), et aussi dans les parages de gènes de développement et de gènes de facteurs transcriptionnels.

La preuve qu'il existe bien un monde totalement méconnu d'ARN vient d'être administrée grâce à la méthode de balayage massif du génome par *tiling array* très dense, assurant une couverture exhaustive du génome. Cinquante millions de sondes ont été utilisées dans la publication de P. Bertone *et al.* [2] qui ont ainsi exploré 1,5 Gb du génome non répété avec 134 *microarrays* de 39 000 sondes chacun, explorant les transcrits émanant des deux brins. En plus des transcrits attendus (ARNm de protéines), ils ont trouvé plus de 10 000 transcrits nouveaux correspondant à des régions non-codantes (intergéniques et introniques). Ils appellent les régions du génome ainsi révélées des

TAR (*transcriptionally active units*). Une seconde publication [3] utilisant le même principe et interrogeant l'ADN de 10 chromosomes humains (environ un tiers du génome) distribué en 70 millions d'oligos (avec un pas de 5 nucléotides) a rapporté des résultats similaires. Dans cet article, les nouveaux transcrits sont appelés TUF (*transcripts of unknown function*), conformément à l'usage en vigueur dans le projet « ENCODE ». Il reste à s'assurer que ces résultats ne sont pas le simple reflet d'un bruit de fond transcriptionnel, c'est-à-dire d'une transcription relâchée quelque peu aléatoire. Si tel n'est pas le cas, cela signifierait que le génome non répétitif, c'est-à-dire environ la moitié du génome total, n'est pas constitué comme on le pensait jusqu'à présent de seulement 2 % d'ADN soi-disant « utiles » parce que codant, les 98 % restants étant qualifiés d'« inutiles » (le « *junk* » DNA) parce qu'on ne sait pas à quoi ils servent. Ainsi, on aurait totalement méconnu jusqu'ici un monde sous-jacent, celui des ARNnc, qui pourrait être 10 fois plus vaste que celui du transcriptome classique. Il reste à en dresser un inventaire et à en analyser les fonctions. Si ce monde est bien réel, ces ARN devraient avoir des fonctions régulatrices, et former un écheciveau qui sera difficile à débrouiller, car ces entités sont souvent transcrites dans les deux sens et chevauchantes, avec des régions de transcription hyperdense (appelées « forêts » de transcription), comme l'a montré une analyse exhaustive de 181 000 transcrits de la souris, dont les limites 5' et 3' ont été caractérisées [4].

Laboratoire de Biochimie et génétique moléculaire, Hôpital et Institut Cochin, 123, boulevard de Port-Royal, 75014 Paris, France.

kaplan@cochin.inserm.fr

Il n'est pas déraisonnable de penser qu'un territoire aussi grand puisse être affecté par des mutations, et que la pathologie jouera, une fois de plus, un rôle d'éclaircisseur. Il existe déjà un certain nombre d'observations de situations pathologiques qui pourraient bien appartenir à cette catégorie [5]. Néanmoins, l'examen attentif de ces publications n'emporte pas toujours la conviction, ce qui laisse présager une validation difficile de la pathogénicité des mutations affectant le nouvel ARNome.

À peine a-t-on déchiffré le génome, qu'un nouveau défi se présente, complètement inattendu et sans doute plus difficile à relever que le simple grand séquençage. Il n'a fallu que 10 ans pour éditer la séquence complète du génome, il faudra beaucoup plus pour la lire.

Neither genes, nor junk, but... TAR/TUF !

RÉFÉRENCES

1. Claverie JM. Fewer genes, more noncoding RNA. *Science* 2005 ; 309 : 1529-30.
2. Bertone P, Stolt V, Royce TE, *et al.* Global identification of human transcribed sequences with genome tiling arrays. *Science* 2004 ; 306 : 2242-6.
3. Cheng J, Kapranov P, Drenkow J, *et al.* Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. *Science* 2005 ; 308 : 1149-54.
4. The FANTOM Consortium, Carninci P, Kasukawa T, Katayama S, *et al.* The transcriptional landscape of the mammalian genome. *Science* 2005 ; 309 : 1559-63.
5. Szymanski M, Barciszewska MZ, Erdmann VA, Barciszewski J. A new frontier for molecular medicine: noncoding RNAs. *Biochim Biophys Acta* 2005 (sous presse).

TIRÉS À PART

J.C. Kaplan

