

Un gène dans un gène...

« *Gene within a gene...* » tel est le titre surprenant d'un article publié par Henikoff *et coll.* dans la revue Cell [1], article démontrant, s'il en était besoin, l'extraordinaire diversité des « solutions » adoptées par la vie. Les auteurs travaillaient initialement sur le gène de drosophile codant pour trois enzymes de la synthèse des purines, dénommé « locus Gart ». Dans des études antérieures, ils avaient démontré que ce gène donnait deux types de transcrits, de 4,7 et 1,7 kilobases (kb), possédant des séquences 5' identiques et différant par leurs extrémités 3', allongées dans le cas du transcrit de 4,7 kb. L'espèce de 4,7 kb est le messenger d'une protéine de 144 000 de poids moléculaire alors que l'espèce de 1,7 kb code pour une protéine de seulement 46 000 daltons. Le mécanisme responsable de la synthèse de ces deux messagers à partir d'un même gène est un épissage alternatif associé, pour produire le messenger de 1,7 kb, à l'utilisation d'un signal de polyadénylation situé dans le

4ème intron (figure 1). Une analyse par ordinateur de la séquence nucléotidique du gène devait révéler que le 1er intron, de 4,12 kb, pouvait coder dans le sens opposé à celui du gène Gart, c'est-à-dire sur l'autre brin d'ADN, pour une protéine de 200 acides aminés.

Afin de déterminer si cette séquence pouvait effectivement être transcrite en un ARN différent de ceux décrits plus haut, un brin complémentaire fut synthétisé chimiquement et utilisé comme sonde pour hybrider avec les ARN de la mouche, dans différents tissus à différents stades de développement. Un ARN de 0,9 kb fut ainsi détecté, mais uniquement dans la larve, alors que les messagers codant pour les enzymes de la synthèse des purines sont présents dans tous les tissus et à tous les stades de développement. La détection du messenger spécifique par hybridation *in situ* sur des coupes histologiques d'embryon permet de confirmer que seules certaines cellules de la larve au stade « pré-pupaire » contenaient le mes-

sager. La protéine synthétisée sous la direction de l'ARN de 0,9 kb a de nombreuses similitudes structurales avec les protéines de la cuticule larvaire et pourrait bien être l'une d'entre elles. Ainsi un même segment d'ADN contient le gène Gart sur un brin, transcrit dans toutes les cellules, et, sur l'autre brin et dans un intron du premier gène, un deuxième gène, codant probablement pour une protéine de la cuticule larvaire spécifiquement synthétisée dans certaines cellules du stade pré-pupaire.

Les questions posées par cette organisation tout à fait étonnante sont nombreuses. La première est ce qui se passe réellement dans les cellules qui expriment les deux gènes à la fois, c'est-à-dire sont transcrits par l'ARN polymérase dans les deux sens. Comment est évitée la « collision » entre ces machineries transcriptionnelles progressant en sens inverse? L'origine de l'imbrication des deux gènes étudiés par Henikoff *et coll.* n'est évidemment pas connue. L'hypothèse à l'heure

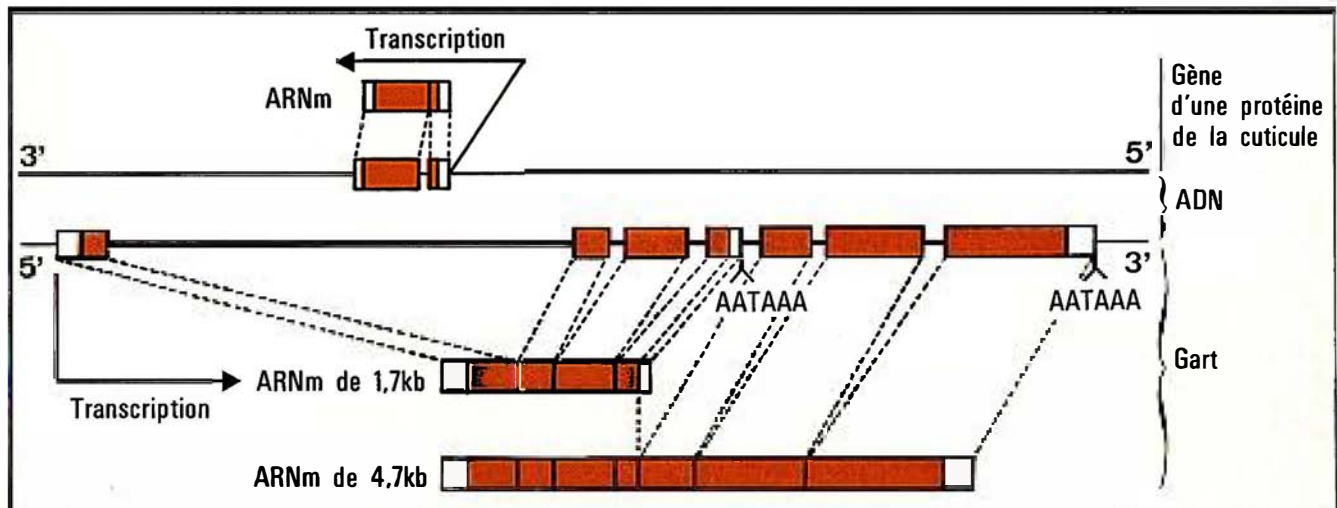


Figure 1. **Organisation et transcription des gènes du locus Gart.** Le brin inférieur de l'ADN code pour les protéines de la synthèse des purines constituant le locus Gart. Lorsque le quatrième intron est excisé, la polyadénylation se fait après le site « AAUAAA » le plus distal, tout à droite du schéma, donnant un messenger de 4,7 kb. En cas de non excision de ce quatrième intron, la polyadénylation se fait après le signal AAUAAA situé dans l'intron, et l'ARN ainsi obtenu n'a que 1,7 kb. Le brin supérieur de l'ADN contient, au niveau de la région correspondant au premier intron du gène Gart, un gène entièrement différent, contenant lui-même un intron et transcrit dans l'autre direction. Ce gène semble coder pour une protéine de la cuticule larvaire de la drosophile. Rectangles rouges : région codante des exons; rectangles blancs : région non codante des exons; traits épais : introns; traits fins : régions intergéniques.

actuelle la plus probable est celle d'une « transposition » du gène codant pour la protéine de la cuticule dans le locus *Gart*, créant ainsi un intron nouveau contenant ce gène. Les phénomènes de « transposition », c'est-à-dire de changement de position de fragments d'ADN transposés d'un site à un autre site du génome, sont d'ailleurs bien connus dans de très nombreuses espèces, notamment les drosophiles.

Une dernière question concerne la fréquence réelle de phénomènes similaires à ceux rapportés ici. Il existe certes d'autres exemples d'introns pouvant coder pour des protéines, principalement au niveau des gènes mitochondriaux de champignons (notamment de levures) et de chloroplastes de plantes. Dans ces cas, cependant, la séquence codante intronique est située sur le même brin que celles des exons du gène dans lequel elle est située, et sa transcription est colinéaire avec celle de ce gène. Les protéines introniques correspondantes pourraient intervenir dans l'épissage [2] ou la transposition [3] de ces introns. L'article commenté ici rapporte donc bien le premier exemple de l'intrication de deux gènes totalement différents dont les contrôles de transcription sont indépendants. Il n'est cependant pas sûr que l'on ait systématiquement recherché des séquences potentiellement codantes sur les deux brins d'introns dont certains sont immenses (jusqu'à 60 kb !) et pourraient bien contenir maintes informations ignorées jusqu'à présent. Nul doute que des dizaines de laboratoires dans le monde ne se soient d'ores et déjà mis à vérifier ce point grâce à leurs ordinateurs; la réponse est donc probablement pour bientôt. A. K.

1. Henikoff S, Keene MA, Fechtel K, Fristom JW. Gene within a gene: nested drosophila genes encode unrelated proteins on opposite DNA strands. *Cell* 1986; 44 : 33-42.
2. Lazowska J, Jacq C, Slonimski, P.P. Sequence of introns and flanking exons in wild type and box 3 mutants of cytochrome b, reveals an interlaced splicing protein coded by an intron. *Cell* 1980; 22 : 333-48.
3. Jacquier A, Dujon B. An intron encoded protein is active in a gene conversion process that spreads an intron into a mitochondrial gene. *Cell* 1985; 4 : 383-94.

Le clonage du facteur VII

Les facteurs de la coagulation sont clonés un à un. Après les facteurs antihémophiliques A et B, c'est le tour du Facteur VII, pour lequel un ADN complémentaire a été obtenu par une équipe de Seattle [1]. Le facteur contient une séquence signal qui, une fois coupée, livre une chaîne de 406 acides aminés dont la séquence a été déduite de celle de l'ADNc. L'intérêt du clonage est d'abord de mieux comprendre le mécanisme d'activation du facteur VII : celui-ci est en effet transformé en facteur VIIa activé par une simple coupure, donnant deux chaînes que maintient réunies un pont disulfure. C'est la chaîne lourde qui porte le site catalytique de cette protéase à sérine, alors que la chaîne légère contient la dizaine de résidus d'acide γ carboxyglutamique qui caractérisent cette protéine vitamine K dépendante. La séquence clonée a permis de placer exactement la coupure, entre une arginine et une isoleucine, et de fixer les

dimensions des chaînes, respectivement de 254 et 152 acides aminés.

Le déficit en facteur VII est rare (de l'ordre de un cas sur un million) et ne justifie pas en soi la préparation de la protéine par génie génétique. Son obtention, possible dans un des systèmes développés pour la biosynthèse des protéines dépendantes de la vitamine K (voir médecine/science n° 6, vol. 1, p. 330) pourrait cependant s'avérer utile pour traiter des hémophiles porteurs d'anticorps anti-facteur VIII [2]. Le facteur VII pourrait en effet, par le biais de la voie extrinsèque de la coagulation, court-circuiter le facteur VIII s'il est présent en quantité suffisante.

J.-C. D.

1. Hagen FS, Gray CL, O'Hara P, et al. Characterization of a cDNA coding for human factor VII. *Proc Natl Acad Sci USA* 1986; 83 : 2412-6.
2. Hedner U, Kisiel W. Use of human factor VIIa in the treatment of two hemophilia A patients with high-titer inhibitors. *J Clin Invest* 1983; 71 : 1836-41.

Une sonde ADNc spécifique du virus de l'hépatite D

L'hépatite D est une maladie souvent redoutable, observée chez des malades porteurs du virus de l'hépatite B [1]; elle est due à un virus, le virus de l'hépatite D (VHD) ou agent delta, incapable de se répliquer dans la cellule en l'absence du virus B, qui est co-infectant [1, 2, 3]. L'enveloppe du virus B (qui porte l'antigène HBs) sert d'enveloppe au virus D. Le génome du virus D est constitué d'une molécule simple brin d'ARN. Afin d'obtenir une sonde pour la détection du virus D, l'ARN viral a été recopié en ADN complémentaire (ADNc) par la transcriptase reverse [4]; la copie ADNc a ensuite été insérée dans un plasmide permettant son amplification. Le clone recombinant obtenu permet de disposer à volonté d'une sonde ADNc. Celle-ci a été utilisée pour détecter le virus dans le sérum de malades infectés. Il suffit de 1 µl de sérum pour obtenir, en hybridation moléculaire sur filtre

(voir médecine/sciences n° 2, vol. 2, p. 104), un bon signal chez un malade à la phase aiguë de l'infection, 0,5 ml étant indispensable pour donner un signal équivalent chez un malade ayant une hépatite chronique [4]. L'ADNc obtenu jusqu'à présent est de petite taille (166 bases) et ne représente que moins du dixième du génome viral (environ 2000 bases). Son isolement constitue néanmoins un premier pas vers la caractérisation totale du génome. La sonde disponible a déjà un intérêt considérable pour le diagnostic et la compréhension de la maladie. S.E.

1. Erlinger S. L'hépatite D. *médecine-sciences* 1985; 1 : 64-5.
2. Bréchet C. L'agent delta : biologie et pathobiologie. *médecine-sciences* 1985; 1 : 66-8.
3. Bernuau J. Les hépatites dues au virus D. *médecine-sciences* 1985; 1 : 69-73.
4. Denniston K J, Hoyer B H, Smedile A, Wells F V, Nelson J, Gerin J L. Cloned fragment of the hepatitis delta virus RNA genome: sequence and diagnostic application. *Science* 1986; 232 : 873-5.