

PRIX NOBEL DE PHYSIOLOGIE ET DE MÉDECINE 1993

NOBEL 93

La nécessité des gènes en morceaux

Phillip A. Sharp et Richard J. Roberts

Phillip A. Sharp, né le 6 juin 1944 à Falmouth (Kentucky, USA), travaille à Cambridge (Massachusetts, USA) au Massachusetts Institute of Technology.

Richard J. Roberts, né le 6 septembre 1943 à Derby (Grande-Bretagne), a travaillé jusqu'en 1992 au laboratoire de Cold Spring Harbor (New York, USA).

Jusqu'en 1977 personne ne se posait le problème de la structure des gènes chez les êtres vivants en général, le dogme étant « ce qui est vrai chez *E. coli* l'est également chez l'éléphant ». Ainsi, personne ne doutait que les gènes eucaryotes fussent identiques à ceux des procaryotes, constitués d'un message continu. C'est dire le coup de tonnerre que furent les publications des laboratoires de Phillip Sharp au *Massachusetts Institute of Technology* et de Richard Roberts à *Cold Spring Harbor Laboratory*. Le feu d'artifice débuta en juillet 1977 par un article de *Cell* [1] dans lequel R.E. Gelin et R.J. Roberts concluaient de l'étude de l'extrémité 5' du messenger tardif de l'adénovirus 2 que « un grand transcrit nucléaire pourrait subir une maturation pour donner naissance à différentes formes d'ARN messagers cytoplasmiques ». En août 1977, c'est dans *Proceedings of National Academy of Sciences of USA* que S.M. Berg et al., du laboratoire de Phillip Sharp, démontraient formellement que l'extrémité 5' terminale du messenger tardif d'adénovirus comportait des éléments variablement épissés. La preuve définitive de cette

hypothèse était apportée par l'étude en microscopie électronique des *R-loops* (boucles R) dans lesquelles un brin de l'ADN du gène majeur tardif hybridé avec ses messagers forme des boucles monobrins au niveau des introns dont il n'existe aucune séquence complémentaire sur l'ARNm [2]. De semblables expériences parfaitement concluantes de boucles R étaient publiées en septembre 1977 par L.T. Chow du laboratoire de R.J. Roberts [3].

Dès novembre 1977, R. Breathnach et al., du laboratoire de Pierre Chambon à Strasbourg, démontraient que le gène de l'ovalbumine de poulet avait, lui aussi, une structure mosaïque, qu'il était transcrit en un ARN précurseur secondairement excisé et épissé pour engendrer l'ARN messenger [4]. Peu après, le même R. Breathnach décrivait, avec ses collègues du laboratoire de Pierre Chambon, la loi générale gouvernant la reconnaissance des jonctions exons/introns (les introns débutent au dinucléotide GT et finissent au dinucléotide AT) [5]. La nature discontinue des gènes des eucaryotes a, dès lors, été très rapidement intégrée comme l'un des phénomènes essentiels du fonctionnement normal et pathologique de la machinerie génétique. Les mécanismes de l'excision/épissage des introns sont d'une extrême complexité et commencent seulement à être connus. Cette complexité explique qu'un défaut de maturation de l'ARN précurseur en messenger soit un mécanisme en cause

dans près de 10 à 20 % des maladies génétiques.

C'est également cette structure discontinue des gènes eucaryotes qui permet aisément la constitution de gènes hybrides fonctionnels en cas de réarrangement chromosomique, tels qu'ils sont très souvent rencontrés dans les cancers. En effet, la reconstitution d'un gène fonctionnel serait bien difficile si sa séquence codante était continue. En revanche, les points de cassure ont beaucoup plus de chance de se situer dans des introns qui représentent la plus grande partie des gènes, aboutissant à un nouveau gène discontinu constitué des exons provenant des deux chromosomes réarrangés. C'est également la structure discontinue des gènes qui permet d'expliquer les mécanismes de création de la diversité des anticorps et des récepteurs de l'antigène des lymphocytes T, secondaires à des réarrangements géniques juxtaposant ou rapprochant des exons différents.

Des introns existent aussi chez des eucaryotes inférieurs, levure et protozoaires. Certains d'entre eux catalysent eux-mêmes leur excision suivie de l'épissage des exons, ce qui fut à l'origine de la découverte des ribozymes, couronnée par le prix Nobel de chimie 1989 décerné à Thomas R. Cech et Sidney Altman (*m/s* n° 9, vol. 5, p. 703). D'autres types d'introns, enfin, possèdent une séquence codante pour des protéines à activité enzymatique intervenant dans le processus d'excision/épissage,

ou bien conférant potentiellement à ces introns des propriétés d'éléments mobiles (voir *Lexique m/s, Gènes, oncogènes et médecine : structure et expression des gènes, 1990, p. 8-10 et [6-8]*). Cette multiplicité des types d'introns a conduit très tôt à poser la question de leur origine (unique ou multiple) et des mécanismes expliquant la présence d'introns chez certains êtres vivants et leur absence chez d'autres. Il est probable que, dans certains cas, les introns sont apparus à l'intérieur des gènes où ils se sont intégrés en tant qu'éléments mobiles. Des introns codant pour une activité enzymatique de type transcriptase inverse/transposase ont probablement été insérés dans certains sites par ce mécanisme [9]. Cependant, l'hypothèse que les premières formes de vie possédaient des introns qui ont été perdus chez les eubactéries actuelles, mais conservés, voire développés chez les eucaryotes, continue d'être rationnelle et attrayante. De nombreuses archéobactéries, supposées être des espèces fossiles, conservent des séquences de type intronique séparant des exons. On peut imaginer que la première molécule d'ARN messager a pu se constituer plus aisément par juxtaposition de séquences oligonucléotidiques discontinues, recopiées sur la matrice génomique originelle

(elle-même composée d'ARN), qu'en identifiant d'emblée sur cette matrice une longue séquence signifiante continue (on parlerait aujourd'hui de « phase ouverte de lecture »). Par la suite, deux types de pression ont pu conduire à la perte ou à la conservation des introns. Lorsque des êtres vivants étaient avant tout sélectionnés sur leur capacité à proliférer rapidement, les introns, qui doivent être répliqués et transcrits, ont pu constituer un fardeau insurmontable en terme de dépense énergétique. En revanche, les êtres pour lesquels la sélection a surtout privilégié la plasticité et la faculté d'adaptation avaient toutes les raisons de conserver leurs introns : le mécanisme de réarrangement des exons aboutissant à la création de nouveaux gènes (phénomènes d'*exon shuffling*) a en effet un extraordinaire potentiel à créer de la diversité à partir d'éléments préexistants. Les phénomènes d'épissage alternatif permettent également de diversifier les conditions d'expression d'un gène, dans différents tissus ou au cours du développement.

Ainsi, la structure discontinue des gènes est-elle très probablement une propriété essentielle du matériel génétique expliquant qu'elle ait été conservée malgré le fardeau énergétique évoqué ci-dessus et l'extraordinaire

complexité des mécanismes nécessaires au rétablissement de la continuité du message.

Axel Kahn
Élisabeth Bursaux

1. Gelinas RE, Roberts RJ. One predominant 5'-undecanucleotide in adenovirus 2 late messenger RNAs. *Cell* 1977 ; 11 : 533-44.
2. Berget SM, Moore C, Sharp PA. Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proc Natl Acad Sci USA* 1977 ; 74 : 3171-5.
3. Chow LT, Gelinas RE, Broker TR, Roberts RJ. An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell* 1977 ; 12 : 1-8.
4. Breathnach R, Mandel JL, Chambon P. Ovalbumin gene is split in chicken DNA. *Nature* 1977 ; 270 : 314-9.
5. Breathnach R, Benoist C, O'Hare K, Gannon F, Chambon P. Ovalbumin gene : evidence for a leader sequence in mRNA and DNA sequences at the exon-intron boundaries. *Biochemistry* 1978 ; 17 : 4853-7.
6. Breathnach R, Chambon P. Organization and expression of eucaryotic split genes coding for proteins. *Annu Rev Biochem* 1981 ; 50 : 349-83.
7. Padgett RA, Grabowski PJ, Konarska MM, Seiler S, Sharp PA. Splicing of messenger RNA precursors. *Annu Rev Biochem* 1986 ; 55 : 1119-50.
8. Abelson J. RNA processing and the intervening sequence problem. *Annu Rev Biochem* 1979 ; 48 : 1035-69.
9. Lambowitz AM. Infectious introns. *Cell* 1989 ; 56 : 323-6.