

Structure des gènes eucaryotes

Les gènes eucaryotes sont constitués de segments codants et non codants d'ADN, appelés respectivement exons et introns. À première vue, il semble que ce soit une charge inutile de transporter l'ADN sans fonctions évidentes dans un gène. Cependant, il a été reconnu que cela a de grands avantages évolutifs. Lorsque des parties de gènes différents sont réarrangés sur de nouveaux sites chromosomiques au cours de l'évolution, de nouveaux gènes peuvent être construits à partir de parties de gènes déjà existants.

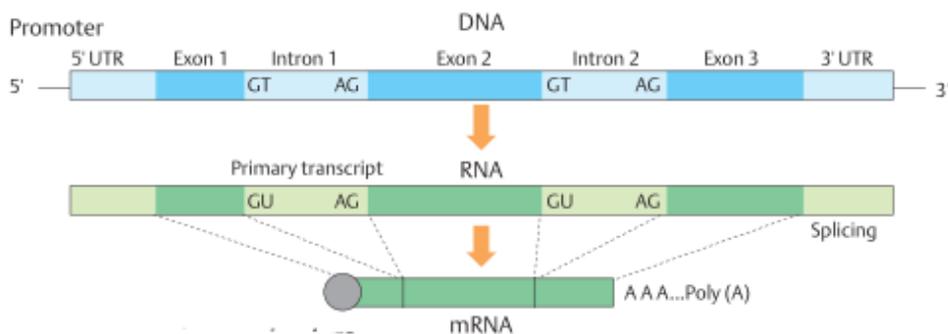
Exons et introns

En 1977, il a été trouvé de manière inattendue que l'ADN d'un gène eucaryote était plus long que son ARNm correspondant. La raison en est que certaines sections du transcrite d'ARN primaire initialement formé sont retirées avant que la traduction ne se produise. Les micrographies électroniques montrent que l'ADN et son transcrite (ARN) ont des longueurs différentes (1). Lorsque l'ARNm et son ADN simple brin complémentaire sont hybrides, des boucles d'ADN monocaténaire se forment parce que l'ARNm s'hybride uniquement avec certaines sections d'ADN simple brin. Dans (2), sept boucles (A à G) et huit sections s'hybridant sont montrées (1 à 7 et la section d'attaque L). Sur les 7700 paires de bases d'ADN de ce gène (3), seulement 1825 s'hybrident avec l'ARNm. Un segment s'hybridant s'appelle un exon. Une section d'ADN initialement transcrite qui est ensuite retirée du transcrite primaire est un intron. La taille et la disposition des exons et des introns sont caractéristiques de chaque gène eucaryote (structure exon / intron). (Micrographie électronique de Watson et al., 1987).

Séquences d'ADN intervenantes (introns)

Chez les procaryotes, l'ADN est colinéaire avec l'ARNm et ne contient aucun intron (1). Chez les eucaryotes, l'ARNm mature n'est complémentaire que de certaines sections de l'ADN, car ce dernier contient des introns (2). (Figure adaptée de Stryer, 1995).

Structure génique eucaryote basique

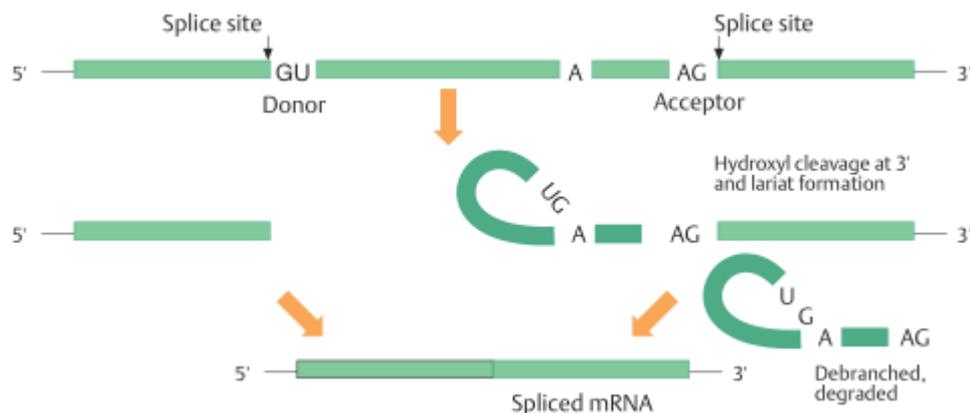


Structure génique eucaryote basique

Les exons et les introns sont numérotés dans la direction 5 'à 3' du brin codant. Les deux exons et les introns sont transcrits en un ARN précurseur (transcription primaire). Le premier et le dernier exons contiennent habituellement des séquences qui ne sont pas traduites. Ceux-ci sont appelés la région 5 'non traduite (5' UTR) de l'exon 1 et la 3 'UTR à l'extrémité 3' du dernier exon. Les segments non codants (introns) sont retirés du transcrit primaire et les exons de chaque côté sont connectés par un processus appelé épissage. L'épissage doit être très précis pour éviter une modification indésirable du cadre de lecture correct. Les introns commencent presque toujours par les nucléotides GT dans le brin 5 'à 3' (GU dans l'ARN) et se terminent par AG. Les séquences à l'extrémité 5 'de l'intron commençant par GT sont appelées site donneur d'épissage et à l'extrémité 3', se

terminant par AG, sont appelées site accepteur d'épissage. ARNm mature est modifié au 5' terminer en ajoutant une structure stabilisatrice appelée «bouchon» et en ajoutant de nombreuses adénines à l'extrémité 3' (polyadénylation).

Voie d'épissage dans les introns



GU-AG

Voie d'épissage dans les introns GU-AG

L'épissage de l'ARN est un processus complexe médié par une grande protéine contenant de l'ARN appelée un splicéosome. Il s'agit de cinq types de petites molécules d'ARN nucléaires (snRNA) et de plus de 50 protéines (petites particules de riboprotéines nucléaires). Le mécanisme de base de l'épissage implique schématiquement un clivage autocatalytique à l'extrémité 5' de l'intron, ce qui entraîne la formation d'un lariat. Ceci est une structure circulaire intermédiaire formée en reliant l'extrémité 5' (UG) à une base (A) à l'intérieur de l'intron. Ce site s'appelle le bras de jonction. Au stade suivant, le clivage au site 3' libère l'intron sous forme de lariat. En même temps l'exon droit est ligaturé (épissé) à l'exon gauche. Le lariat est déramifié pour donner un intron linéaire et ceci est rapidement dégradé. Le site de la branche identifie l'extrémité 3' pour un clivage précis au site accepteur d'épissage. Il se trouve à 18-40 nucleotides en amont (dans la direction 5') du site d'épissage 3'. (Figure adaptée de Strachan et Read, 1999).