

dans lesquelles il évolue, les perturbations causées par un micro-organisme pathogène... Après avoir décrit, à travers quelques exemples, les ARN mixtes, leurs fonctions et leur importance dans certaines pathologies, nous envisagerons les mécanismes évolutifs ayant permis l'éclosion de ces ARN qui, tels des couteaux suisses, permettent aux organismes les exprimant de faire face à une multitude de situations dans un environnement en perpétuel changement.

ARN messagers et régulateurs, une frontière floue

Le premier ARN non codant identifié a été l'ARN de transfert alanine, purifié à partir de levures, dont la structure a été publiée en 1965 par le biochimiste américain Robert Holley – travaux récompensés par un prix Nobel en 1968. Cet ARN est impliqué dans la synthèse des protéines : il transporte, jusqu'à la protéine en construction, un constituant des protéines (l'acide aminé alanine) et l'insère dans la chaîne protéique aux endroits codés par l'information génétique. Depuis cette découverte, grâce aux progrès des techniques de séquençage, des centaines d'ARN non codants ont été identifiés dans les génomes de procaryotes (cellules sans noyau) et des milliers chez les eucaryotes (avec noyau). Leurs fonctions sont peu à peu élucidées. La majorité d'entre eux semblent réguler l'expression de gènes cibles en intervenant à diverses étapes du processus conduisant des gènes à la synthèse des protéines. Leur taille varie de quelques nucléotides (les constituants de l'ADN et de l'ARN) à plusieurs dizaines de milliers.

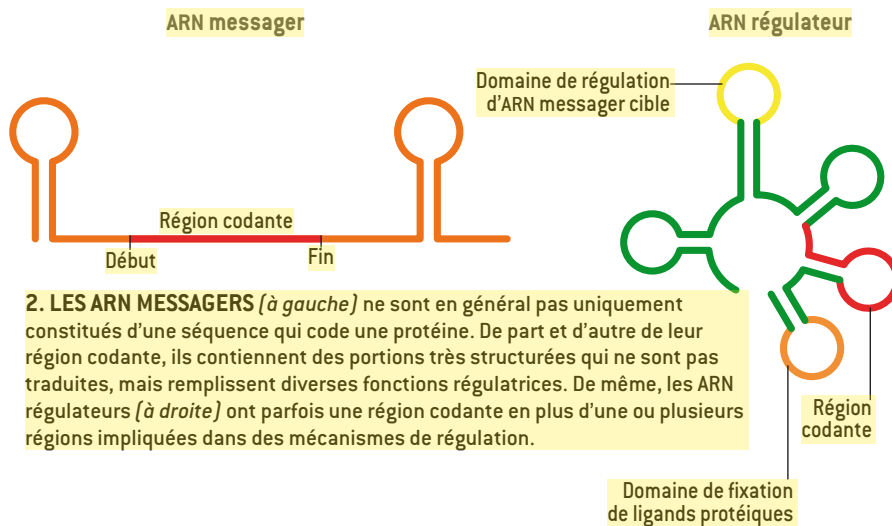
Au cours de la dernière décennie, des progrès techniques spectaculaires ont permis de séquencer le patrimoine génétique de nombreux organismes vivants, des virus à l'homme. Désormais, du fait de la baisse des coûts de séquençage, le nombre de génomes séquencés d'individus distincts au sein d'une même espèce augmente à grande vitesse. Cela permet d'analyser et de comparer les génomes d'entités biologiques similaires provenant de zones géographiques différentes. Une fois les

séquences génomiques déterminées, des approches bio-informatiques automatisées différencient le matériel génétique contenant de l'information codante de celui non traduit en protéines, à l'aide de signatures spécifiques présentes au sein de la séquence même des génomes.

Chez les organismes procaryotes, l'information codante est majoritaire dans le génome. Chez les eucaryotes, dont nous faisons partie, c'est l'inverse : la majeure partie du génome n'encode pas de protéines, mais contient une information colossale sous forme d'une panoplie de gènes transcrits en ARN fonctionnels de différentes tailles et fonctions. Seul un pour cent environ des génomes de mammifères contient de l'information traduite en protéines, alors que 70 à 90 pour cent sont transcrits au cours de la vie en un transcriptome vaste et complexe d'ARN non codants. Le catalogue de ces transcrits fait l'objet d'une recherche internationale active. Il est encore incomplet, mais permet déjà de dresser un nouveau portrait des ARN.

Les transcriptomes des organismes vivants contiennent un mélange d'ARN de tailles et de fonctions variées, dont les ARN messagers ne sont qu'un sous-ensemble. Un ARN messager se définit par la séquence codante qu'il contient, construite par une succession de triplets de nucléotides – des codons –, dont chacun spécifie un acide aminé, avec un triplet marquant le début et un autre la fin de la séquence codante. Ces triplets sont traduits par les ribosomes, les usines de transformation de l'information génétique en protéines, avec l'aide des ARN de transfert qui les approvisionnent en acides aminés (voir l'encadré pages 54 et 55). Généralement, en amont et en aval de leurs séquences codantes, les ARN messagers contiennent des portions transcrites, mais non traduites, très structurées, dont

1. LES ARN SONT LES MOLÉCULES TRANSCRITES à partir de l'ADN. On découvre aujourd'hui que nombre d'entre eux remplissent plusieurs fonctions. Un même ARN présente plusieurs domaines plus ou moins structurés ayant chacun une fonction différente, symbolisée par une couleur sur cet ARN fictif.



2. LES ARN MESSAGERS (à gauche) ne sont en général pas uniquement constitués d'une séquence qui code une protéine. De part et d'autre de leur région codante, ils contiennent des portions très structurées qui ne sont pas traduites, mais remplissent diverses fonctions régulatrices. De même, les ARN régulateurs (à droite) ont parfois une région codante en plus d'une ou plusieurs régions impliquées dans des mécanismes de régulation.

certaines ont des fonctions régulatrices (voir la figure 2). Par divers mécanismes, ces régions modulent la transcription de l'ARN messenger, sa stabilité ou sa traduction en protéine.

Un ARN régulateur présente quant à lui des propriétés fonctionnelles liées à sa séquence et à son repliement (son organisation spatiale), qui lui permettent d'interagir avec une ou plusieurs macromolécules. Certains ARN régulateurs portent aussi des séquences qui sont traduites par des ribosomes en chaînes protéiques – des peptides – dans des conditions physiologiques spécifiques. Comme les ARN messagers, ces séquences sont constituées d'une succession de codons, chacun spécifiant un acide aminé, les premier et dernier codons étant des signaux d'initiation et de terminaison de la traduction, reconnus par les ribosomes et leurs partenaires. Les fonctions des peptides encodés par les ARN régulateurs sont liées, ou non, aux actions régulatrices des ARN qui les contiennent.

Des ARN mixtes dans tous les organismes

Ainsi, nombre d'ARN messagers agissent en tant qu'ARN régulateurs et, à l'inverse, on découvre de plus en plus d'ARN régulateurs portant des séquences codantes. Une séparation stricte entre ces deux classes de macromolécules apparaît d'autant moins réaliste que, chez les eucaryotes, certains ARN non codants, dits longs, ont une taille comparable à celle des ARN messagers, supérieure à 200 nucléotides. Or chez différents mammifères, des souris jusqu'à l'homme, beaucoup de ces ARN non codants longs contiennent des séquences codantes qui sont exprimées dans des contextes spécifiques (au

cours du développement ou dans un type cellulaire particulier) ou alors très faiblement. Comment, dans ce cas, les différencier d'ARN messagers ? C'est pour éviter cette difficulté que nous proposons de nommer ARN mixtes les ARN multifonctionnels.

Détectés dans l'ensemble des organismes (voir la figure 3), les ARN mixtes ont au minimum deux fonctions, l'une liée à la chaîne protéique exprimée par l'ARN, l'autre produite par une interaction spécifique de l'ARN avec un partenaire, entraînant la régulation de gènes cibles. Les mécanismes mis en œuvre sont variés : interaction avec un ARN messenger, qui bloque sa traduction ou, au contraire, l'active, guidage d'une enzyme vers un ARN messenger, qui sera alors coupé, imitation d'une séquence ADN qui leurre la machinerie de traduction, séquestration de protéines impliquées dans l'expression des gènes, catalyse enzymatique (voir l'encadré pages 54 et 55)...

L'ARN de certains virus, par exemple, contient bien plus que de l'information génétique traduite en protéines fonctionnelles. À leur extrémité non traduite, ces ARN viraux portent des éléments essentiels à leur réplication. Parfois, leurs extrémités ressemblent à des ARN de transfert. De tels ARN ont ainsi au moins trois fonctions distinctes : ils encodent de l'information génétique, contiennent des éléments essentiels à la copie de leur génome et sont des pseudo-ARN de transfert.

Dans des cas plus subtils encore, l'ARN régule positivement certaines cibles, négativement d'autres et exprime en plus un peptide fonctionnel ! C'est le cas de l'ARNIII exprimé par les staphylocoques dorés, des bactéries responsables de nombreuses infections, notamment nosocomiales, et de l'ARN *SgrS* produit par deux autres bactéries, *Escherichia coli* et la salmonelle.

L'ARNIII intervient dans la régulation de l'expression coordonnée de gènes par l'ensemble d'une population bactérienne. Dans le mécanisme qui coordonne cette expression à l'échelle d'une population, nommé « détection du quorum », l'ARNIII sert de déclencheur : en fonction de la densité bactérienne, il régule l'expression de gènes permettant aux bactéries de passer de la phase d'adhérence aux cellules de l'hôte à leur colonisation. Il produit aussi un peptide – une hémolysine – qui détruit les globules rouges (voir l'encadré page 56).

L'ARN *SgrS*, quant à lui, est impliqué dans le maintien de l'équilibre en glucose

INDEX

ARN
(ou acide ribonucléique).
Molécule transcrite à partir du génome (l'ADN).

ARN messagers
Ensemble des ARN qui codent des protéines.

ARN régulateurs non codants
Ensemble des ARN qui ne codent pas de protéines. Ils ont diverses fonctions régulatrices dans la cellule.

ARN mixtes
ARN agissant à la fois comme messenger et comme régulateur.

ARN antisens
ARN capables d'interagir spécifiquement, par appariements avec un ou plusieurs gènes cibles.

Transcriptome
Ensemble des ARN exprimés à la suite de la transcription d'un génome par les enzymes ARN polymérases.