

UNE DÉCOUVERTE RÉCENTE : L'INTERFÉRENCE DE L'ARN (d'après rapport AGREG 2007 & <http://www.snv.jussieu.fr/vie/dossiers/siRNA/index.htm>)

C'est en travaillant sur les mécanismes cellulaires et moléculaires de la coloration des fleurs et en particulier du pétunia qu'une observation étonnante fut faite dans les années 1990. La couleur mauve des pétunias est produite par l'action d'une enzyme, la chalcone synthase. On savait que la perte de fonction du gène qui code cette synthase provoque une perte plus ou moins marquée de la coloration de la fleur. Afin d'obtenir des pétunias plus colorés, des copies supplémentaires du gène ont été insérées dans le génome de la plante par transgénèse. Les plantes obtenues avaient toutes des fleurs blanches. L'analyse des ARN isolés à partir des fleurs blanches a montré que le taux d'ARNm produit par le gène était fortement réduit par rapport au taux d'ARN des fleurs contrôles. Phénotypes extrêmement troublants, ***tout se passait comme si les gènes introduits «éteignaient» la production de l'enzyme.*** Les gènes ajoutés et l'endogène sont rendus «silencieux». D'autres expériences montrent que l'introduction d'ARN double-brin dans des cellules d'*Arabidopsis thaliana* déclenche une méthylation de l'ADN correspondant. Cette méthylation inhibe la transcription du gène donc la production de l'ARNm correspondant.

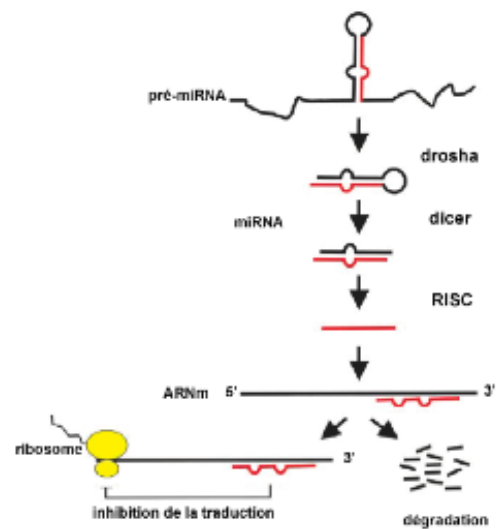
À la fin des années 1990, pour chercher à comprendre la fonction du gène *par1* chez le ver nématode *Caenorhabditis elegans*, l'ARN complémentaire de l'ARNm de ce gène fut injecté dans l'embryon. Cette technique classique permet de générer des ARN double-brin dans les cellules puisque l'ARN complémentaire injecté va s'apparier avec l'ARNm endogène. Le duplex formé empêchera la traduction et donc la protéine correspondante ne sera pas produite. Parmi les contrôles de ce type d'expérience, il faut envisager l'injection d'un excès d'ARNm ou l'injection de l'ARNm hybridé à son brin complémentaire. Dans le premier cas, on s'attend à une augmentation de la production de la protéine, dans l'autre à une production normale. Il n'en fut rien. Avec ces deux types de contrôle, il fut montré que la protéine n'était pas produite.

Le principe général du mécanisme mis en évidence par ces expériences est maintenant en partie élucidé. Il s'agit de la mise en œuvre de la **machinerie d'interférence de l'ARN**.

Les microARN

Le clonage et l'analyse systématique de gènes impliqués dans des mutations chez le ver nématode *Caenorhabditis elegans*, a contribué à isoler deux gènes particuliers. Ces gènes ne codent pour aucun produit protéique mais pour des ARN d'une longueur de 20-22 nucléotides. Les analyses ont montré qu'ils s'hybrident aux régions 3' non traduites d'ARNm et que les protéines correspondantes ne sont pas produites. Cette découverte a conduit à la recherche systématique des petits ARN. Le séquençage de banques de petits ARN ou la prédiction de séquences par différents algorithmes bio-informatiques a permis d'identifier différents ARN de petite taille, chez les plantes et les animaux. Cette nouvelle classe d'ARN a été nommée **microARN (miARN)**.

Les miARN sont codés par plusieurs centaines de gènes (environ 400 chez l'homme) et sont donc produits de façon naturelle par la cellule. On sait qu'ils sont synthétisés sous forme de pré-miARN d'environ 70 nucléotides. Ils se replient pour former des structures partiellement double-brin présentant une tige boucle en «épingle à cheveux» avec des mésappariements dans la tige de l'épingle. Ils sont maturés dans le noyau par le complexe protéique Drosha. Les miARN matures sont exportés du noyau par l'exportine 5. Dans le cytoplasme, ils sont pris en charge par la protéine Dicer qui produit des ARN de 21-24 bases. Puis, la machinerie d'interférence de l'ARN fonctionne et si la complémentarité entre le miARN et l'extrémité 3' de l'ARNm cible est parfaite, l'ARNm cible est alors dégradé. Au contraire si la complémentarité est imparfaite alors la traduction du message est bloquée.



Le mécanisme de l'inhibition transcriptionnelle par les ARN interférents

Les miARN ou microARN et les siARN ou ARN interférents ne sont pas codés par la même partie du génome. Les premiers (ARNmi) proviennent de gènes qui leur sont propres et dont ils sont les uniques produits. Les seconds (ARNsi) dérivent de deux sources, exogène ou endogène. La source exogène correspond soit à l'apport extérieur d'un ARN double brin par injection expérimentale, par exemple dans le nématode *C. elegans*, soit à un apport en provenance de génomes à ARN double brin comme ceux de certains virus à ARN des plantes. La source endogène correspond à une production naturelle par utilisation des parties exoniques (codantes ou non codantes) d'un gène plus étendu.

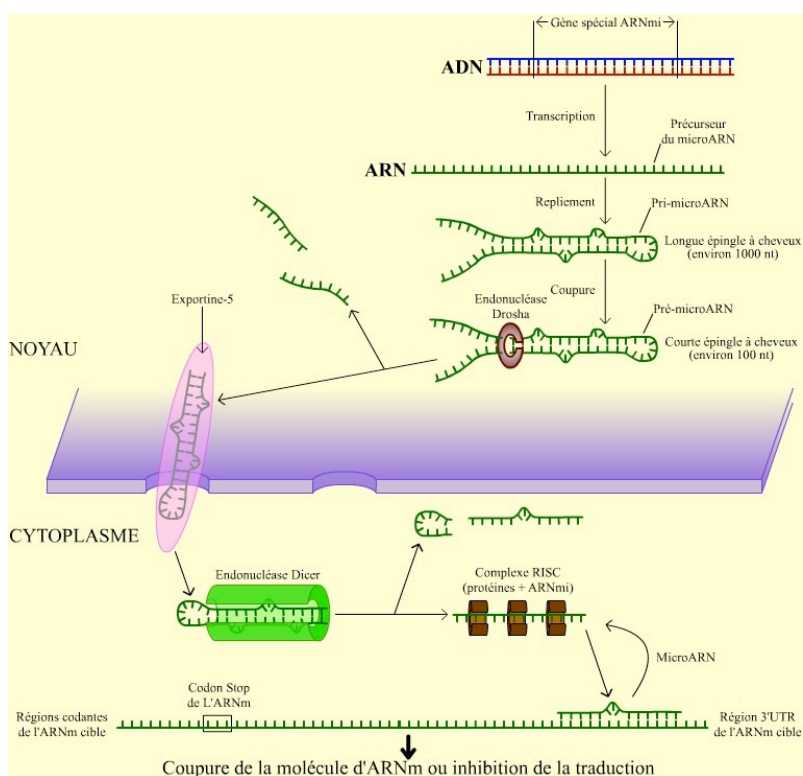
Pour les ARNm, leur transcrite est d'abord une molécule d'ARN simple brin d'un millier de nucléotides. Chacun de ces ARNs présente le long de sa séquence des motifs nucléotidiques intrachânes partiellement complémentaires qui les conduisent à adopter une structure secondaire en forme d'une longue épingle à cheveux. Cette structure peut être divisée en trois parties : le "corps" qui est formé de séquences plus ou moins complémentaires présentes sur le simple brin, se disposant en vis-à-vis par appariement, la "tête" (en forme de boucle) et les "jambes" simple brin de l'épingle qui portent les séquences n'ayant pas trouvé de complémentaires. Se produit alors la série de transformations suivante :

L'enzyme **Drosha** (ribonucléase de type III) y effectue une première série de coupures qui réduit cette structure en une courte molécule d'une centaine de nucléotides, toujours en épingle à cheveux (short hairpin RNA : shRNA), mais amputée de ses "jambes". L'exportine-5 la transporte alors dans le cytoplasme.

Une seconde ribonucléase, appelée **Dicer**, achève le travail commencé par l'enzyme Drosha : par une deuxième série de coupures, elle élimine la "tête" de l'épingle, dernier réduit des séquences non appariées, tandis que d'autres protéines qui lui sont associées (de type hélicase) dissocient le petit ARNd subsistant, d'une vingtaine de nucléotides, en deux ARN simple brin. L'enzyme Dicer clive et sépare également les longs ARN double brin des ARNs en morceaux d'environ 21 nucléotides.

L'un des simples brins obtenus est le microARN (ARNmi) ou le petit ARN interférent (ARNsi). Il s'associe à un complexe protéique en formant un nouveau complexe appelé **RISC** (RNA Induced Silencing Complex). C'est en cet équipage qu'il se porte sur sa cible : l'ARNm auquel il s'apparie.

En cas d'appariement parfait (ARNsi), l'ARNm est détruit et il n'y a pas traduction. Et les RISC qui sont à l'origine de cette destruction restent ensuite parfaitement fonctionnels, ce qui leur permet d'opérer de nouveau sur d'autres ARNm de même spécificité. C'est cette réutilisation qui les rend particulièrement offensifs. Mais même lorsque l'appariement est imparfait (ARNmi), il y a tout de même barrage à la traduction de l'ARNm. A travers l'exemple des ARNm, la figure ci-dessous résume le mode d'action des ces petits ARN.



Les petits ARN : outils de défense naturelle et de surveillance

La machinerie d'interférence à ARN est un mécanisme de défense des cellules contre les virus et les éléments génétiquement mobiles. Les éléments génétiquement mobiles, les **transposons**, sont des séquences d'ADN qui se dupliquent et s'insèrent n'importe où dans le génome. Chez les plantes, il existe des mutations des gènes codant la protéine Dicer ou les protéines du complexe RISC. Les plantes mutées manifestent une augmentation du déplacement des transposons dans leurs génomes. Elles ont également une sensibilité accrue vis-à-vis des infections par les virus à ARN. Ceci s'explique par le fait que les ARN double-brin produits lors de la répllication des virus à ARN sont reconnus par la protéine Dicer. Il en résulte l'activation de la machinerie d'interférence à ARN qui dégrade les ARNm viraux.

La machinerie d'interférence à ARN est également un mécanisme de surveillance de la production d'ARN dans les cellules. En effet, chez les plantes et *Caenorhabditis elegans*, il existe des ARN polymérases ARN dépendantes. Lors de la production des ARNm ou lors de certains dysfonctionnements cellulaires, des transcrits incomplets ou codant des protéines indésirables sont produits. Cette polymérase peut reconnaître ces ARN et produire des ARN double-brin complémentaires des transcrits indésirables. Ces ARN sont émincés par l'enzyme Dicer, et la machinerie d'interférence est activée. Le complexe RISC scanne les ARNm cellulaires. S'il y a une complémentarité entre les petits ARN et un ARNm cible, le complexe RISC le coupe et le dégrade rapidement. La synthèse de la protéine correspondante est impossible. Les petits ARN assurent ainsi un processus de surveillance en réduisant au silence les gènes qui menacent le bon fonctionnement cellulaire.

Chez beaucoup d'organismes à l'exception des mammifères, les petits ARN interférents peuvent être multipliés, amplifiés car ils possèdent l'ARN polymérase ARN dépendante. Ces petits ARN interférents amplifiés peuvent être dispersés dans la lignée germinale et par conséquent passer de génération en génération avec la protection naturelle qui leur est associée. Ils peuvent également diffuser dans tout l'organisme pour conférer leur protection. C'est en particulier le cas chez les végétaux où il a été montré que le transfert des petits ARN interférents a lieu à travers les plasmodesmes. Ce mécanisme confère à la plante entière une résistance à l'infection virale.

Ainsi, l'interférence de l'ARN est une machinerie de lutte pour rendre inactif le matériel génétique des agents infectieux, parasites, virus ou transposons qui s'introduisent dans les cellules ou dans l'ADN. C'est un **véritable système immunitaire intracellulaire**. Il protège le code génétique des espèces.

Les microARN : régulateurs du développement embryonnaire

Dans la plupart des cas, un ARNm peut lier plusieurs miARN et réciproquement un miARN peut se lier à différents ARNm. Cette particularité en fait des régulateurs extrêmement puissants du fonctionnement des cellules. Généralement, les miRNA sont impliqués dans la régulation de facteurs de transcription, indispensables au développement. Chez les nématodes, ils déterminent le passage d'un stade larvaire à l'autre durant le développement post-embryonnaire. Chez la drosophile, ils contrôlent la prolifération et/ou la différenciation cellulaires en modulant l'expression des gènes pro-apoptotiques (apoptose = suicide cellulaire ou mort cellulaire programmée).

Chez l'Homme, leur rôle, pas encore élucidé, est probablement majeur, puisque les gènes codants les miARN ont été conservés au cours de l'évolution.

Chez les plantes, tous les organes latéraux de la partie aérienne, feuilles, pièces florales (pétales ou étamines), sont issus des méristèmes. Ces structures contiennent les cellules souches de la plante. Une organogenèse continue fait qu'au sein des méristèmes, un groupe de cellules va s'individualiser, bourgeonner en un primordium et engendrer tel ou tel organe. Primordia et méristèmes sont séparés entre eux par une frontière formée par un ruban de cellules. Chez *Arabidopsis thaliana*, ces cellules expriment les gènes *CUC1*, 2 et 3. Ils codent pour des facteurs de transcription, comprenant chez *Arabidopsis* plus d'une centaine de membres. L'inactivation d'un ou de plusieurs d'entre eux conduit à des fusions entre les organes ce qui suggère des défauts de spécification des frontières. Par exemple, le double mutant *cuc1 cuc2* se caractérise par une fusion des cotylédons, leur donnant une apparence de coupe et une fusion entre sépales et entre étamines. L'analyse des séquences en 3' des ARNm a révélé que les transcrits des gènes *CUC1* et *CUC2* possèdent un site de fixation pour un miARN, ce site est absent du transcrite de *CUC3*. La surexpression de ce miARN provoque une diminution du niveau d'accumulation des transcrits des gènes *CUC1* et 2 qui conduit à des anomalies embryonnaires ou florales similaires à celles des doubles mutants. Des mutations dans les gènes *CUC1* et *CUC2* ont été générées au niveau du site de fixation du miARN pour empêcher la formation de duplex miARN/ARNm. Les plantes exprimant la forme mutée de *CUC1* produisent un nombre anormal d'organes floraux, une augmentation du nombre de pétales.

Les plantes exprimant la forme mutée de *CUC2* montrent quant à elles un élargissement des frontières. Ces données suggèrent que le miARN régule le développement des frontières entre primordia et méristème. Elles suggèrent également que les miARN régulent l'expression des gènes impliqués dans le développement floral.

La cellule utilise donc les miARN pour contrôler l'expression de son génome, notamment au cours du développement embryonnaire. Ils peuvent éteindre de façon coordonnée des ensembles de gènes dans telle ou telle cellule afin d'acquies une fonction particulière ou l'engager dans une voie de différenciation.

Les ARN interférents : outils pour étudier la fonction d'un gène et pour la thérapie.

Depuis longtemps différentes techniques destinées à étudier la fonction d'un gène sont disponibles. Elles consistent à sur-exprimer une protéine en introduisant l'ARNm qui la code dans la cellule ou l'organisme. Elles consistent également à faire exprimer par la cellule ou l'organisme une protéine mutée. Elles consistent enfin à inhiber la production d'une protéine en empêchant la traduction de son ARNm ou en favorisant sa dégradation. Les outils les plus connus sont les **ARN antisens**, les **ribozymes**, les **oligonucléotides antisens**, ou les **morpholinos**. On analyse ensuite les perturbations qui résultent de ces interventions. Elles permettent d'en déduire la fonction de la protéine. Par rapport à toutes ces techniques, l'ARN interférent semble être à la fois plus efficace et techniquement plus simple à mettre en œuvre.

On peut transfecter des bactéries avec des ARN double-brin capables de détruire l'ARNm d'un gène donné. Ces bactéries sont données en nourriture aux vers. Il en résulte que les ARN double-brin infectent pratiquement toutes les cellules où ils vont inhiber la production du produit d'expression du gène considéré. En observant le phénotype résultant de l'interférence, on peut en déduire la fonction du gène. Puisque les séquences des gènes sont assez bien conservées dans le règne animal, on peut ainsi obtenir des informations transposables aux vertébrés.

Chez l'homme, dans de nombreuses pathologies, certains gènes sont sur exprimés ou exprimés au mauvais endroit ou au mauvais moment. La possibilité de pouvoir inhiber ces expressions pathologiques par les ARN interférents suscite de grands espoirs pour soigner ces maladies. Cependant certains problèmes se posent. En particulier, il faut s'assurer que les ARN interférents gagnent les cellules cibles. Il faut vérifier qu'ils n'introduisent pas de toxicité, être certain de leur stabilité et donc des doses à administrer. Enfin, il faut s'assurer de leur spécificité. Une limitation vient aussi de ce que, contrairement aux plantes, les mammifères ne possèdent pas le système enzymatique qui permet l'amplification des ARN interférents. Cela limite leur disponibilité dans la cellule.

Conclusions

Dans les cellules eucaryotes, les ARN remplissent au moins **trois rôles distincts et complémentaires**.

En premier lieu l'ARN est le **support temporaire de l'information génétique**. C'est l'ARNm qui est utilisé par la cellule pour transmettre l'information correspondant à un gène donné à l'extérieur du noyau, puis pour synthétiser des protéines à partir de ces informations. Il est l'intermédiaire entre le plan directeur stocké dans l'ADN et les agents d'exécution que sont les protéines.

En second lieu, les ARN sont des **catalyseurs enzymatiques**. Comme les protéines, les ARN peuvent se replier pour former des structures complexes. Ces structures permettent à certains ARN de se comporter comme des enzymes, on parle alors de **ribozyme**. Le ribosome, la ribonucléase P et certains introns sont des ribozymes. La machinerie d'épissage des ARNm, le spliceosome est également un ribozyme.

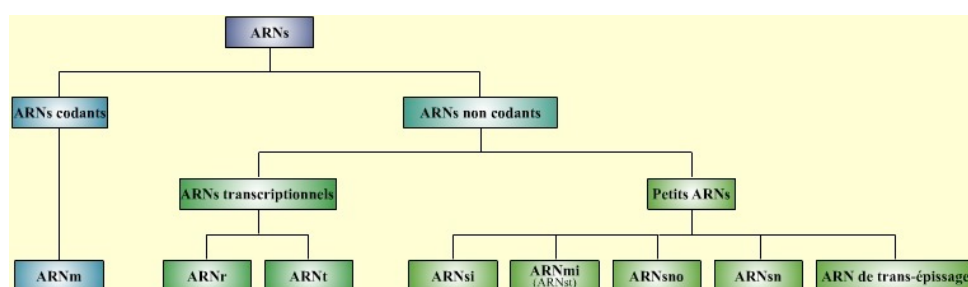
Enfin, les ARN servent de **guide pour des enzymes**. Certains ARN sont utilisés comme co-facteurs par des protéines pour permettre leur ciblage vers des séquences spécifiques. Parmi ceux-ci, on peut citer les petits ARN nucléolaires, qui guident les enzymes de modification de l'ARN ribosomique, ou encore les **ARN interférents**.

Les ARN peuvent **réguler l'expression des gènes** et s'avèrent capables de moduler et de diversifier l'information codée par les gènes. Ils apparaissent comme des acteurs du fonctionnement des cellules et du développement des organismes vivants, au même titre que les protéines qui ont été longtemps considérées comme les éléments fondamentaux de la régulation du fonctionnement cellulaire. Ils sont codés par des séquences d'ADN distinctes des séquences codant pour les protéines. Cela permet d'élargir la définition même du gène. On peut considérer **qu'un gène est une séquence d'ADN à l'origine d'une molécule fonctionnelle, que ce soit une protéine ou un ARN**.

L'ARN peut être considéré comme une molécule très polyvalente, ce qui a conduit à proposer une hypothèse selon laquelle l'ARN serait la plus ancienne de toutes les macromolécules biologiques. C'est la théorie, dite du **monde à ARN**. Cette théorie propose que l'ARN, capable de combiner à la fois des fonctions de catalyseurs et d'information génétique serait le précurseur universel. L'ARN serait antérieur à l'ADN comme support de l'information génétique, ce qui expliquerait ses fonctions plus étendues. L'ADN serait apparu plus tard et n'aurait supplanté l'ARN que pour le rôle de stockage de l'information à long terme.

Un des arguments qui va dans ce sens est la **plus grande stabilité de l'ADN**. Cette stabilité est conférée par la disparition d'une fonction alcool au niveau du ribose. La stabilité de l'information génétique est aussi liée au fait que **l'ADN est bicaténaire**. Cet aspect est extrêmement important car, si une mutation se produit sur un brin de l'ADN, la cellule peut corriger cette erreur en se basant sur la complémentarité des brins en utilisant la machinerie de réparation dont elle dispose, ce qui n'est pas possible avec l'ARN.

Nous avons vu que le dogme central du fonctionnement du vivant repose sur la trilogie ADN – ARNm – protéine. Nous avons vu qu'il existe dans les cellules eucaryotes une grande quantité d'ARN non codants provenant de la transcription du génome. Ces ARN jouent un rôle important dans le contrôle de l'activité des gènes d'une manière séquence-spécifique. Ils remettent en cause le dogme fondateur de la biologie auquel il faut introduire un **contrôle épigénétique exercés par les ARN**, qui peuvent assurer protection et réarrangement des génomes.



La famille des ARNs

ARNm : ARN messenger
 ARNr : ARN ribosomique
 ARNt : ARN de transfert
 ARNsi : small interfering RNA ou petit ARN interférent
 ARNm (ARNst) : micro ARN (qui comprennent les ARNst pour small temporal RNA ou petit ARN temporaire)
 ARNsno : small nucleolar RNA ou petit ARN nucléolaire
 ARNsn : small nuclear RNA ou petit ARN nucléaire