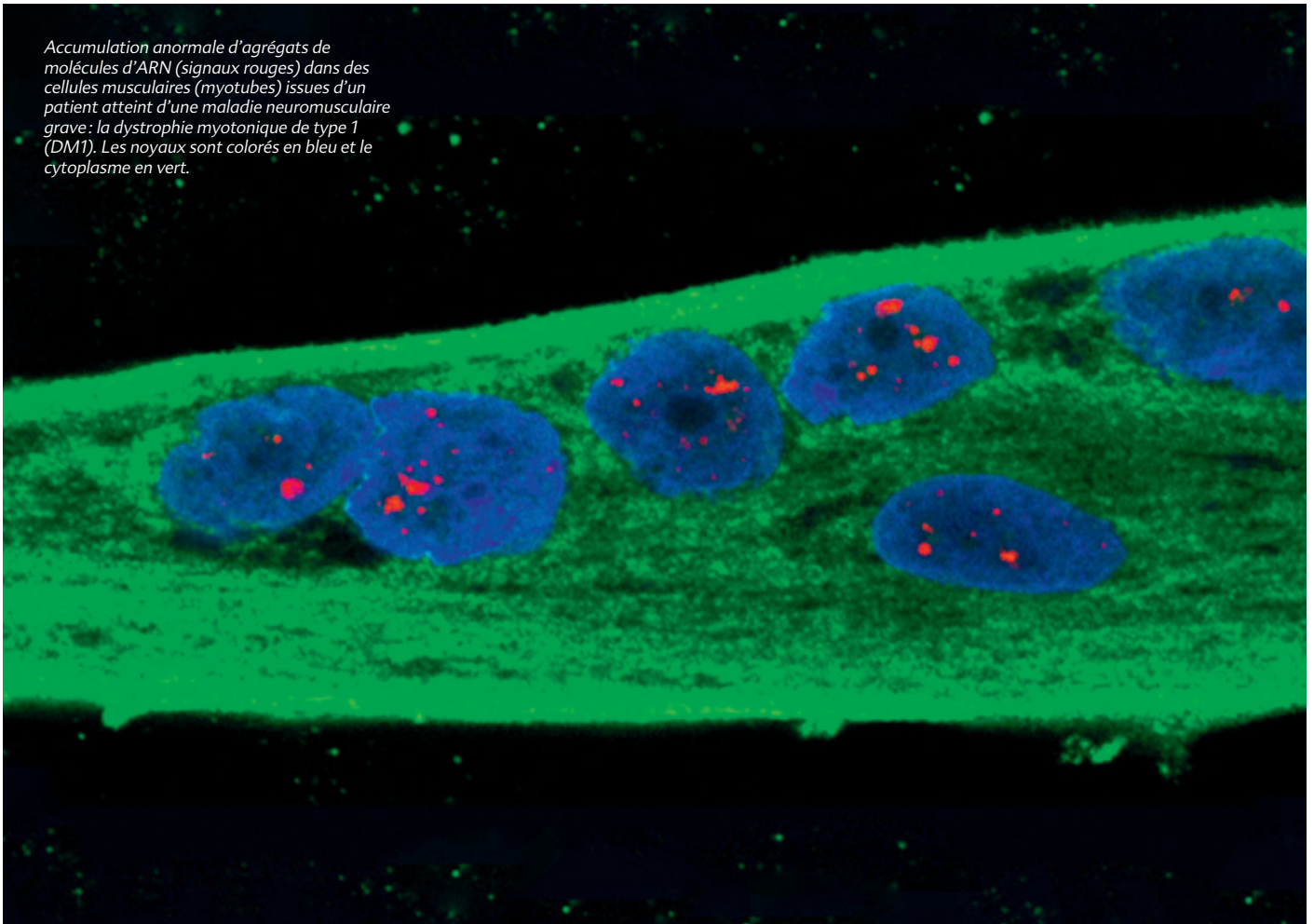


Accumulation anormale d'agrégats de molécules d'ARN (signaux rouges) dans des cellules musculaires (myotubes) issues d'un patient atteint d'une maladie neuromusculaire grave: la dystrophie myotonique de type 1 (DM1). Les noyaux sont colorés en bleu et le cytoplasme en vert.



L'ARN au cœur de la vie

ARN. Trois lettres qui désignent l'Acide RiboNucléique, une classe de macromolécules biologiques apparentées à l'ADN. Capables de porter de l'information génétique, les ARN peuvent également intervenir directement dans des réactions biologiques. Avec l'explosion de la connaissance des génomes, le monde des ARN est actuellement en pleine expansion et son étude représente une des thématiques de recherche les plus fécondes de « la révolution post-génomique ».

Depuis le milieu du XX^e siècle, on sait que l'information génétique est transmise de génération en génération par l'ADN, un polymère biologique très stable qui compose les chromosomes dans notre génome. En quelques décennies, l'évolution des technologies de séquençage a permis d'accéder à la lecture complète des génomes dans de très nombreux organismes, à commencer par l'enchaînement des 3,2 milliards de nucléotides qui composent les chromosomes chez l'homme. Un des enjeux majeurs de la biologie moderne consiste à décrypter les mécanismes qui régulent

l'expression génique: comment l'information génétique est-elle utilisée de manière différentielle dans les tissus en formation d'un embryon, dans une bactérie pathogène, dans les différents tissus d'un organisme, dans une cellule cancéreuse dans une plante d'intérêt agronomique? C'est au cœur de cette problématique qu'apparaît l'ARN, un maillon central de l'expression génique.

Durée de vie limitée

L'information portée par l'ADN ne peut pas être directement utilisée par les machineries biosynthétiques de nos cellules. La lecture d'un segment

d'information génétique, autrement dit un gène, commence toujours par sa copie ou transcription sous forme d'ARN, des molécules composées de nucléotides comme l'ADN. Capables de diffuser dans la cellule, chimiquement réactives, d'une durée de vie limitée et régulée, les ARN constituent la forme dynamique de l'information génétique. Certains d'entre eux servent de matrice pour la synthèse des protéines en fournissant le code génétique qui spécifient l'enchaînement des acides aminés. Mais loin d'être de simples copies, ces « ARN messagers » (ARNm) sont les substrats de modifications qui amplifient la

Le big bang de l'ARN

diversité de l'information génétique. Il existe en effet de nombreux mécanismes post-transcriptionnels capables de modifier les ARNm parmi lesquels l'excision de segments de nucléotides ou le changement ciblé de l'identité de certains nucléotides. Par ailleurs, la régulation de la stabilité des ARNm par des mécanismes de surveillance ou leur transport, voire leur confinement, dans certains compartiments subcellulaires sont des mécanismes centraux dans le contrôle temporel et spatial de l'information génétique. L'ensemble de ces modifications — complexes et combinatoires — engendre une telle diversité que selon les tissus, les groupes de cellules ou les stades de développement, un gène est susceptible de produire des dizaines d'ARNm différents capables de dicter la synthèse d'autant de protéines aux fonctions variées et parfois opposées.

98 % du génome inexploré

À côté des ARNm, qui ne correspondent qu'à une faible part du génome des organismes complexes (moins de 2 % de l'ADN chez l'Homme), on trouve les ARN non-codants (ARNnc).

La complexité et la diversité du répertoire de ces molécules, qui ne portent pas d'information codant une protéine, égalent — et peut-être dépassent — celles des ARNm. Certains ARNnc sont minuscules, tels les microARN constitués d'une vingtaine de nucléotides, alors que d'autres comportent jusqu'à plusieurs centaines de milliers de nucléotides... Les ARNnc exercent

ces diverses fonctions sous la forme d'assemblages moléculaires stables, comprenant le plus souvent une molécule d'ARN associée à un ensemble spécifique de protéines. L'ARN peut alors servir d'échafaudage sur lequel viennent s'agencer des protéines dont l'activité ou la réactivité est ainsi modifiée. Ils jouent également le rôle de senseurs moléculaires captant et intégrant des signaux provenant de l'environnement, ou encore de guide permettant l'appariement du complexe sur une cible. Certains ARNnc se comportent même comme de véritables enzymes et catalysent, tout comme des protéines, des réactions chimiques, à l'image des ribosomes qui synthétisent la formation des chaînes polypeptidiques. Le répertoire des ARN non-codants ne cesse de s'étendre grâce à de nouvelles techniques de séquençage massif du contenu des cellules en ARN, dévoilant peu à peu les secrets des 98 % du génome encore inexploré. L'importance fonctionnelle de la grande majorité d'entre eux demeure incertaine mais ils semblent principalement fonctionner comme des modulateurs du flux de l'expression des gènes suivant des modalités très diverses: maintien de l'intégrité et de la stabilité des génomes au fil des générations; régulation de la transcription des gènes; contrôle de la synthèse des protéines; altération de la localisation et l'activité de certaines protéines. À la lumière de l'importance du contrôle post-transcriptionnel et du nombre toujours croissant de fonctions originales assurées par ARNnc,

les travaux de recherche sur les ARN ont des répercussions directes dans le domaine de la santé. De nombreuses études décrivent des altérations qui aboutissent à des anomalies dans la production - et donc la fonction - des ARNnc ou des ARNm. De tels dysfonctionnements apparaissent parfois liés causalement à l'apparition de certains cancers et autres pathologies humaines. À ce titre, certains ARN sont considérés aujourd'hui non seulement comme de nouveaux bio-marqueurs à valeur diagnostique (ou pronostique), mais aussi représentent très certainement de nouvelles cibles thérapeutiques lorsqu'il s'agira de corriger ou d'atténuer les fonctions défectueuses.

Point fort de la biologie toulousaine

La recherche sur les mécanismes de l'expression génique en général, et la biologie de l'ARN en particulier, est historiquement un des points forts de la biologie toulousaine. Elle est déclinée sur des organismes-modèles très différents, pour des études fondamentales, biomédicales ou agronomiques et bénéficient des plateformes de séquençage et d'analyse des génomes regroupées au sein de la Génopole Toulouse Midi-Pyrénées. Les articles de ce dossier en proposent quelques exemples à travers différents laboratoires. ■

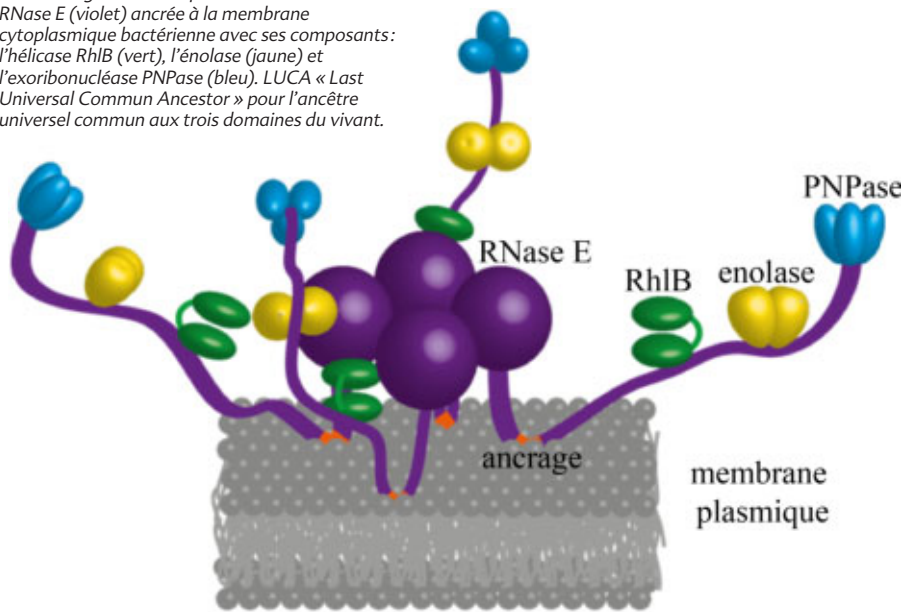


Contact

pierre-emmanuel.gleizes@ibcg.biotoul.fr
et Jerome.cavaill@ibcg.biotoul.fr

Jérôme Cavaillé, directeur de recherche CNRS et **Pierre-Emmanuel Gleizes**, professeur UPS, chercheurs au Laboratoire de biologie moléculaire eucaryote (LBME, unité mixte UPS/CNRS)

Echafaudage tétramérique de l'endoribonucléase RNase E (violet) ancrée à la membrane cytoplasmique bactérienne avec ses composants: l'hélicase RhlB (vert), l'énolase (jaune) et l'exoribonucléase PNPase (bleu). LUCA « Last Universal Common Ancestor » pour l'ancêtre universel commun aux trois domaines du vivant.



Béatrice Clouet-d'Orval, chargée de recherche CNRS, **Marie Bouvier**, chargée de recherche CNRS, et **Agamemnon James Carpousis**, directeur de recherche CNRS, tous chercheurs au Laboratoire de microbiologie et de génétique moléculaires (LMGM, unité mixte UPS/CNRS)



Contact

Beatrice.clouet-dorval@ibcg.biotoul.fr

Vie et mort des ARN

Les **ribonucléases** régulent l'expression des gènes en contrôlant la maturation des ARN, mais aussi leur dégradation. Elles s'associent parfois au sein de complexes moléculaires aux compétences multiples et pourraient avoir été présentes aux tous débuts de l'histoire de la vie.

L'expression des gènes et sa régulation conditionnent le développement de chaque cellule vivante. La cascade de mécanismes est initiée par la transcription de l'ADN en ARN messager (ARNm) et se termine par la traduction de l'ARNm en protéine. La biogénèse et le devenir des ARN sont façonnés par des mécanismes moléculaires complexes. Leur défaillance aboutit à la production d'ARNm erronés ou à une déficience en ARN essentiels (ARN ribosomiques, ARN de transfert, ARN régulateurs) et conduit à des phénotypes mutants chez les bactéries et des pathologies chez l'humain.

Les ribonucléases sont des acteurs clés au cœur de dispositifs macromoléculaires complexes permettant la dégradation du polymère d'ARN. Ce sont en général des enzymes protéiques ayant une activité catalytique permettant soit de fragmenter l'ARN (cas des endoribonucléases), soit d'en grignoter les extrémités (cas des exoribonucléases). À ce jour, plus de 30 familles de ribonucléases ont été répertoriées dans les trois domaines du vivant. Ce nombre reflète la diversité des fonctions à accomplir par ces enzymes dans la cellule. Ce sont des facteurs essentiels de la maturation des molécules d'ARN à partir des longs précurseurs, des composants intrinsèques des systèmes du contrôle-qualité des ARN

défectueux, des gardiens contre les invasions de génomes viraux et de la propagation d'éléments transposables, et enfin des effecteurs des voies de signalisation qui contrôlent l'expression des gènes au niveau post-transcriptionnel. La connexion de ces mécanismes aux autres voies essentielles de la cellule permet de concevoir un modèle intégré de la régulation de l'expression génique.

Machinerie

L'activité de certaines ribonucléases est coordonnée au sein d'une grande machinerie de dégradation des ARN à multiples composants. C'est le cas pour la ribonucléase RNase E, protéine essentielle chez la bactérie *Escherichia coli* et conservée chez les entérobactéries, qui s'associe à une hélicase à ARN, l'énolase et une exoribonucléase pour former une machinerie multi-protéique, nommée dégradosome à ARN. Cette machinerie a un rôle clé dans la dégradation des ARNm, la maturation d'ARN essentiels et la régulation post-transcriptionnelle des gènes par des ARN régulateurs. La RNase E étant ancrée à la membrane cytoplasmique, les machineries de transcription et de dégradation se trouvent séparées. L'objectif est maintenant de comprendre la pertinence physiologique de cette localisation.

Origines évolutives

On se demande également comment un système aussi visiblement élaboré est apparu au cours de l'évolution. La biologie au XX^e siècle a reclassé les organismes modernes en trois domaines primordiaux: eucaryotes, bactéries et archées. Les archées, pour la plupart extrêmophiles, sont des micro-organismes modèles d'intérêt fondamental pour l'étude des mécanismes moléculaires conservés et pour le décryptage de l'histoire évolutive de la vie. Il est donc intéressant d'identifier les enzymes et complexes multi-protéiques de dégradation ou de maturation des ARN chez les archées, et de présager leurs origines évolutives. L'exploration du métabolisme des ARN dans ce domaine du vivant est devenue possible grâce au développement d'outils bio-informatiques performants et au nombre croissant de génomes disponibles. Les ribonucléases de la famille β -CASP en sont l'illustration. Leur présence dans la grande majorité des génomes bactériens et dans tous les génomes eucaryotes et archéens suggère une fonction essentielle et fondamentale des ribonucléases β -CASP dans l'expression génique et leur existence dans la cellule ancestrale commune aux trois domaines du vivant. ■

Le grand meccano du ribosome

Les ribosomes sont des machines moléculaires essentielles et universelles, constituées d'ARN et de protéines, qui servent de plateforme pour la synthèse des protéines. L'assemblage de ces complexes macromoléculaires constitue un meccano géant au coeur du métabolisme cellulaire.

Les ribosomes sont présents dans l'ensemble du monde vivant (à l'exception des virus qui utilisent les ribosomes de leur hôte) et assurent dans chaque cellule l'un des plus anciens mécanismes de la vie: la synthèse des protéines. Scannant le code génétique porté par les ARN messagers, ils catalysent la formation des liaisons entre les acides aminés qui, une fois assemblés forment les protéines naissantes. Les ribosomes sont un exemple d'ARN possédant des propriétés catalytiques, un mécanisme mieux connu depuis l'obtention par cristallographie rayons X de la structure atomique de ribosomes bactériens. Cette découverte valut le prix Nobel de chimie à Venkatraman Ramakrishnan, Tom Steitz et Ada Yonath en 2009.

Un processus à haut débit

Chez l'homme, 4 ARN et plus de 80 protéines s'assemblent pour former les deux sous-unités du ribosome, qui constituent une particule d'environ 30 nm de diamètre. Produits à raison de plusieurs dizaines de milliers d'unité par cellule et par heure, les ribosomes représentent en moyenne 80% des ARN cellulaires. Leur

formation est donc un processus à haut débit surveillé par de nombreuses étapes de « contrôle qualité ». Sa dérégulation est associée à l'émergence de tumeurs cancéreuses, ce qu'exploitent de nouvelles thérapies anti-cancéreuses ciblant cette voie biosynthétique.

Ribosomopathies

Les sous-unités ribosomiques résultent d'un processus complexe d'assemblage et de maturation au cours duquel plus de 150 protéines accessoires et une centaine de petits ARN (snoARN) interviennent pour cliver les chaînes d'ARN ribosomique, en modifier les nucléotides, chaperonner leur repliement ou réguler leur association avec les protéines constitutives. Comprendre les mécanismes moléculaires de la maturation des particules pré-ribosomiques est un thème de recherche majeur de deux équipes au LBME. Elles centrent actuellement leurs études sur le rôle des snoARN, des hélicases (enzymes capable d'ouvrir des hélices d'ARN), ou encore des protéines ribosomiques. Elles ont aussi démontré l'impact de mutations entraînant des altérations de la synthèse des ribosomes dans des maladies



Pierre-Emmanuel Gleizes, professeur UPS, **Annie Mougin**, chargée de recherche CNRS, **Yves Henry**, directeur de recherche CNRS, **Clément Joret**, doctorant, **Natasha Larburu**, doctorante, **Célia Plisson-Chastang**, chargée de recherche CNRS, **Kamila Belhabich-Baumas**, ingénieure d'étude CNRS, **Marlène Faubladier**, ingénieure de recherche UPS, **Jean-Paul Gélugne**, maître de conférences UPS, **Marie-Françoise O'Donohue**, chargée de recherche CNRS, **Régine Capeyrou**, ingénieure de recherche CNRS, **Anthony Henras**, chargé de recherche CNRS, tous au Laboratoire de biologie moléculaire des eucaryotes (LBME, unité mixte UPS/CNRS)



Contact

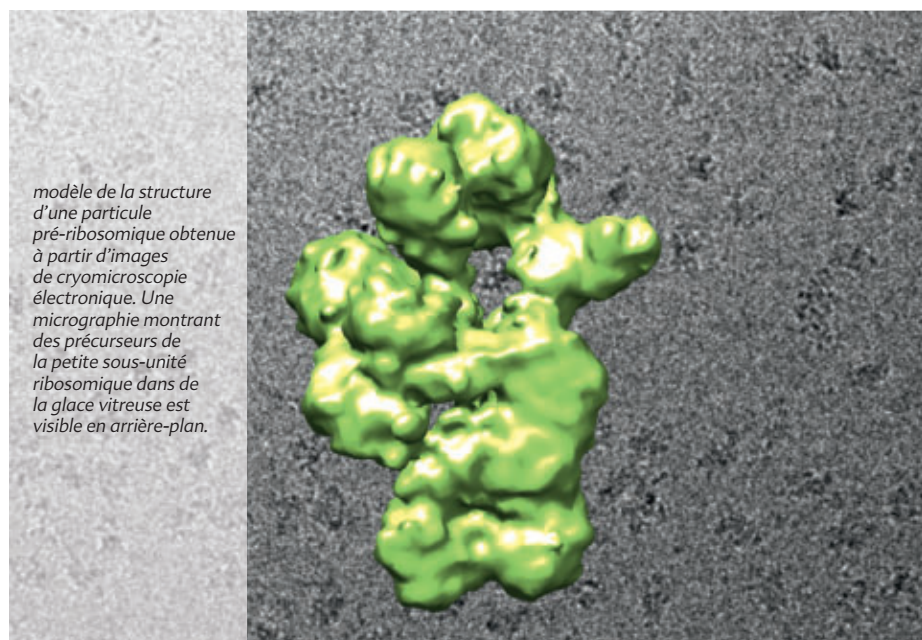
gleizes@ibcg.biotoul.fr,
henry@ibcg.biotoul.fr

génétiques rares comme l'anémie de Blackfan-Diamond ou la dyskératose congénitale, au point qu'elles sont désormais regroupées sous le vocable « ribosomopathies ».

Structure tridimensionnelle reconstruite

Un des défis majeurs du domaine consiste à comprendre comment ces dizaines d'acteurs moléculaires interagissent entre eux à l'échelle de la structure des précurseurs ribosomiques. Dans ce but, les équipes du LBME combinent les approches moléculaires et génétiques à la cryomicroscopie électronique. Cette méthodologie, permet d'observer du matériel biologique dans un état hydraté à l'échelle subnanométrique. Le caractère transitoire des différents précurseurs ribosomiques complique leur étude structurale. La levure, *Saccharomyces cerevisiae*, dont on peut enrichir des particules pré-ribosomiques à certaines étapes de maturation fournit un excellent modèle. Après extraction et purification, les particules sont congelées ultra-rapidement dans un film de glace vitreuse et visualisées par microscopie électronique à transmission. La structure tridimensionnelle est alors reconstruite grâce à l'analyse statistique de milliers d'images. La reconstruction de la structure de particules précurseurs à des étapes successives permettra d'accéder à une nouvelle « vision » des mécanismes de maturation des ribosomes. ■

Ces travaux ont bénéficié en particulier du soutien de la Ligue contre le cancer.



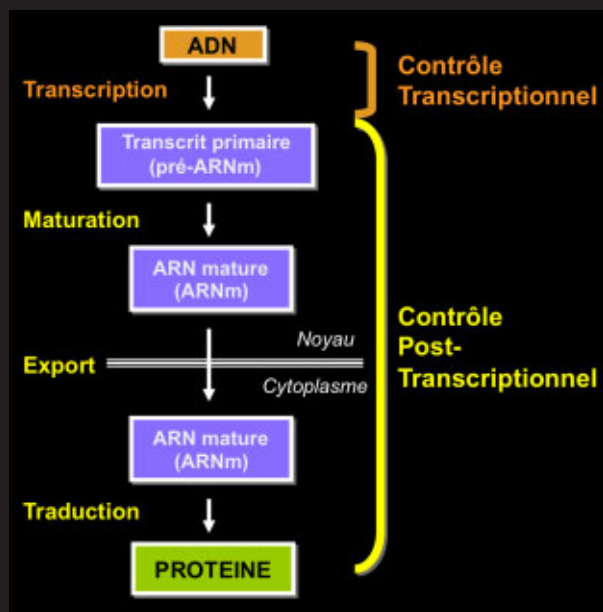
modèle de la structure d'une particule pré-ribosomique obtenue à partir d'images de cryomicroscopie électronique. Une micrographie montrant des précurseurs de la petite sous-unité ribosomique dans de la glace vitreuse est visible en arrière-plan.

De l'ADN aux protéines: les messagers sous contrôle

Les **ARN messagers** portent l'information génétique qui détermine la synthèse des protéines. La modification des ARNm, de même que le contrôle de leur durée de vie ou de leur localisation, sont des niveaux de régulation essentiels de l'expression des gènes.



Hervé Prats, professeur UPS, **Stefania Millevoi**, chargée de recherche Inserm et **Stéphane Pyronnet**, directeur de recherche Inserm, chercheurs au Centre de recherches de cancérologie de Toulouse (CRCT, unité mixte Inserm/UPS)



L'expression génique est un processus en plusieurs étapes. L'information biologique stockée sous forme d'ADN est d'abord transcrite en pré-ARNm. Après maturation dans le noyau, l'ARNm est exporté dans le cytoplasme où il est traduit en protéine. L'expression génique est contrôlée par des mécanismes transcriptionnels, mais aussi post-transcriptionnels.



Contact

herve.prats@inserm.fr
stefania.millevoi@inserm.fr
 et stephane.pyronnet@inserm.fr

Transformation en cellule cancéreuse

Une des équipes du CRCT étudie les mécanismes moléculaires contrôlant la traduction des ARNm matures en protéines et la manière dont ces mécanismes sont régulés voire altérés par divers signaux ou stress. Elle a identifié chez des patients des altérations facilitant la traduction d'ARNm en protéines impliquées dans le processus de transformation d'une cellule normale en cellule cancéreuse. En partenariat avec l'industrie du médicament et les hôpitaux de Toulouse, elle tente de développer des molécules potentiellement toxiques uniquement pour les cellules cancéreuses en ciblant sélectivement ces altérations.

Une autre équipe du CRCT s'intéresse aux différents mécanismes de régulation post-transcriptionnelle impliqués dans l'expression de protéines clés dans le développement tumoral. C'est le cas du contrôle de la maturation en 3' du messager codant la protéine p53, acteur majeur du développement cancéreux puisque ce gène est muté dans plus de la moitié des cas de cancer. Les travaux de l'équipe ont récemment montré que l'étape nucléaire de maturation du pré-ARNm contribue à la reprogrammation de l'expression de p53. Nos études se focalisent maintenant sur l'identification des acteurs moléculaires qui interviennent dans l'activation ou l'inactivation de p53. Par ailleurs, l'équipe explore une étape précise du contrôle traductionnel de gènes impliqués dans l'angiogenèse, mécanisme essentiel imposé par la tumeur qui prolifère. La compréhension fine des régulations traductionnelles auxquelles sont soumis ces gènes est cruciale pour comprendre l'action de drogues visant à bloquer le développement mais aussi la dissémination métastatique de ces tumeurs. ■

L'expression génique correspond de manière générale au passage de l'ADN, qui stocke l'information biologique, à la protéine qui réalise la fonction biologique. Chez les eucaryotes (dont les animaux et les plantes) elle suit différentes étapes. L'ADN est tout d'abord transcrit dans le noyau en ARN messager (ARNm) primaire. Chaque extrémité des ARNm primaires est modifiée par des motifs biochimiques particuliers: la « coiffe » de l'extrémité 5' (où commence la synthèse) et la queue de poly-adénine à l'extrémité 3' (où la synthèse se termine). Certains segments des ARNm primaires sont excisés (épissage des introns). Les ARNm matures sont alors exportés du noyau vers le cytoplasme où ils sont traduits en protéines.

Régulation post-transcriptionnelle

Jusqu'aux années quatre-vingt, l'idée prévalait que la régulation de l'expression génique s'opérait principalement à l'étape de sa transcription des gènes en ARNm. Depuis, d'autres niveaux de

régulation essentiels, dits post-transcriptionnels, ont été découverts. En effet, les étapes multiples du processus de maturation des ARNm offrent une grande diversité de modes de contrôle qui sont nécessaires à la spécialisation des cellules (qui va permettre par exemple aux cellules de la rétine de fabriquer une protéine détectant la lumière du soleil...) ou à leur adaptation à différents signaux et stress environnementaux (comme une protéine induite par les UV du soleil brunit la peau...).

Ainsi, lorsque les cellules sont exposées à un stress, certains ARNm codant des protéines impliquées dans la réponse au stress, sont préférentiellement contrôlés. En court-circuitant la transcription, la régulation post-transcriptionnelle permet une réponse plus rapide, précise et adaptée. De plus, la dérégulation des étapes post-transcriptionnelles est également associée à de nombreuses pathologies, dont les cancers. Le rôle de ces mécanismes dans le maintien de l'homéostasie cellulaire s'avère primordial.

ARN: Petits mais ils font le maximum...

Moins de 2 % du génome humain participe directement à la synthèse protéique. Est-ce à dire que la très grande partie de notre ADN n'a pas de réelles fonctions? Certainement pas! Les estimations les plus récentes montrent qu'un accroissement de la part « non-codante » des génomes est corrélé avec la complexité des organismes biologiques qui les hébergent. C'est cette partie obscure, et encore mystérieuse, de notre patrimoine génétique que nos équipes scrutent, avec une attention particulière portée sur les petits ARN non-codants ou ARN régulateurs.

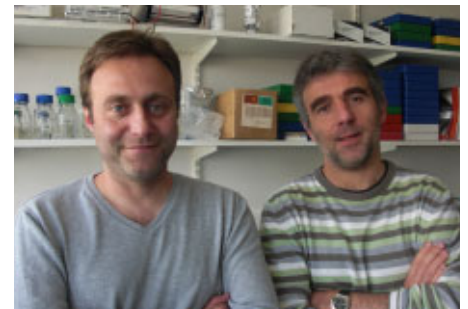
Il apparaît aujourd'hui que certains de ces ARN non-codants - peut-être tous - représentent une nouvelle couche de régulation de l'expression des gènes. Parmi eux, les petits ARN régulateurs sont actuellement sur le devant de la scène et leur étude constitue une thématique forte et dynamique du LBME. On y trouve les petits ARNC/D requis pour la modification chimique d'autres ARN, les microARN impliqués dans le blocage de la synthèse des protéines, mais aussi le petit ARN 7SK qui est un régulateur de la transcription des gènes. En combinant recherches bio-informatiques et séquençages à haut débit, des équipes du LBME ont récemment identifié des centaines de petits ARN dont l'expression, les fonctions présumées et la faible conservation à travers les espèces sont tout à fait originales et inattendues.

Séquences poubelles

C'est par exemple le cas des ARN AluACA, présents dans le noyau des cellules humaines. Fait remarquable, ces ARN sont issus d'éléments mobiles très abondants: les séquences Alu dont le million de copies est dispersé dans le génome des primates, y compris au sein de gènes codant pour des protéines. La communauté scientifique a longtemps considéré de telles séquences répétées comme relativement inutiles, et dont les transcrits ARN étaient probablement éliminés. Nos équipes ont démontré qu'une partie des séquences Alu n'était pas dégradée mais donnait naissance à de petits ARN stables: les AluACA.

Un défi aux lois de l'hérédité!

Pour la grande majorité, nos gènes utilisent les deux copies parentales. Une faible fraction,



Patrice Vitali, maître de conférences UPS et Jérôme Cavaillé, directeur de recherche CNRS, chercheurs au Laboratoire de biologie moléculaire eucaryote (LBME, unité mixte UPS/CNRS).



Contact

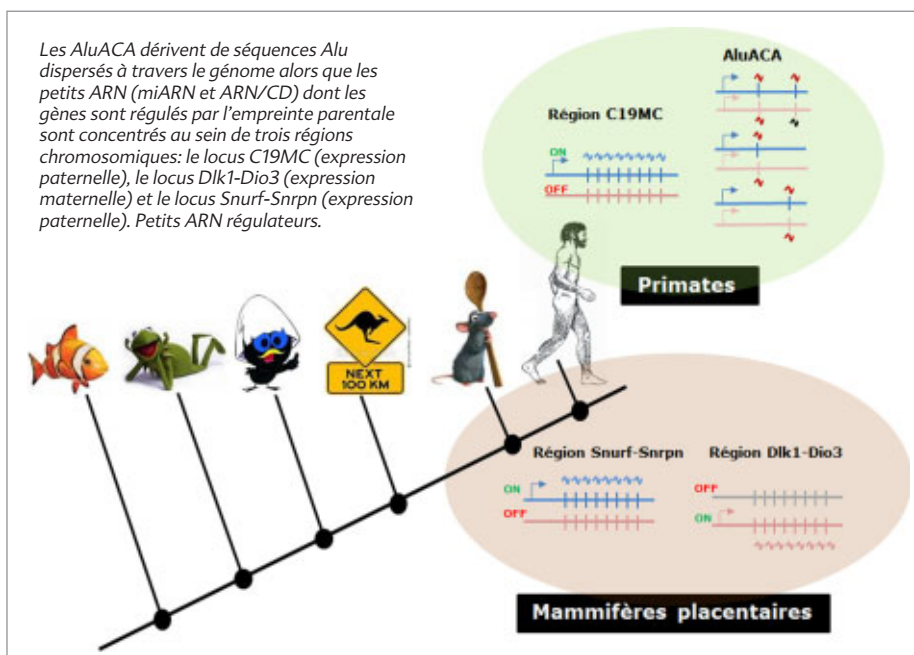
cavaillé@biotoul.fr
vitali@ibcg.biotoul.fr

soumise à l'empreinte parentale n'utilise qu'une copie. Pour un gène donné, cela peut-être la copie d'origine paternelle, pour un autre celle transmise par la mère. Il existe donc « une mémoire » capable de distinguer et de modifier l'activité des deux copies. En cela, l'empreinte parentale défie les lois classiques de l'hérédité mendélienne. Fait remarquable, de nombreux microARN et petits ARN C/D, tous regroupés au sein de trois régions chromosomiques, sont soumis à ce type de phénomène. Ces ARN sont relativement peu conservés au sein des espèces. Certains sont présents exclusivement chez les primates, alors que d'autres sont retrouvés uniquement chez les mammifères placentaires.

Bestiaire

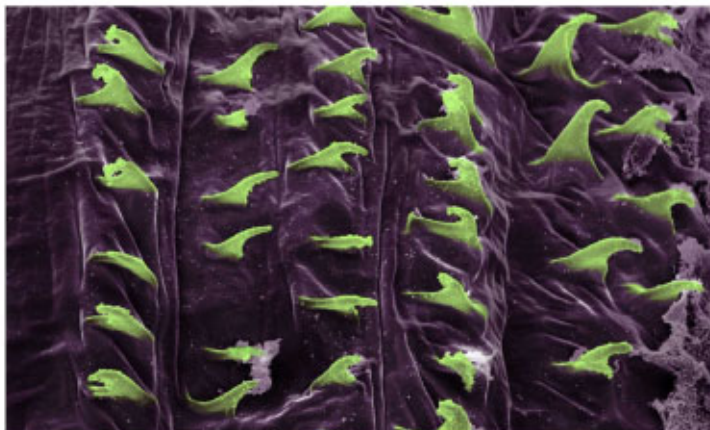
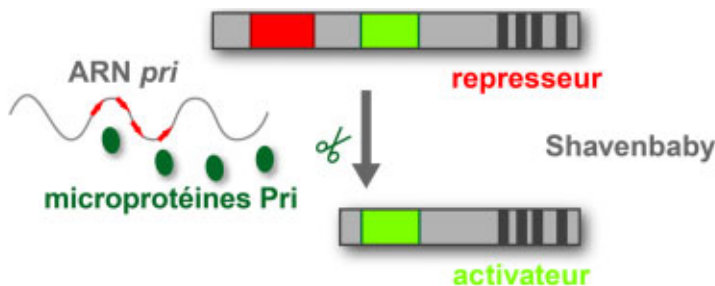
Les études en cours au LBME visent à décrypter les fonctions moléculaires et physiologiques de ces nouveaux petits ARN, y compris via l'étude de modèles murins génétiquement modifiés. Ces recherches ont un réel impact dans le domaine biomédical puisque des altérations dans l'expression des gènes soumis à l'empreinte parentale sont fréquemment décrites dans des cancers et sont responsables de maladies humaines. C'est notamment le cas du syndrome de Prader-Willi dans lequel on constate des anomalies de l'expression d'un petit ARN C/D du locus Snurf-Snrpn. Les quelques exemples cités dans cet article, loin d'être exhaustifs, illustrent la nécessité d'explorer ce véritable « bestiaire à petits ARN » dont l'étendu du répertoire et l'impact sur l'organisation, la fonction et l'évolution des génomes de mammifères en général, et des primates en particulier, demeurent encore largement énigmatiques. ■

Les AluACA dérivent de séquences Alu dispersés à travers le génome alors que les petits ARN (miARN et ARN/CD) dont les gènes sont régulés par l'empreinte parentale sont concentrés au sein de trois régions chromosomiques: le locus C19MC (expression paternelle), le locus Dlk1-Dio3 (expression maternelle) et le locus Snurf-Snrpn (expression paternelle). Petits ARN régulateurs.



De ARNs supposés non codants, sources de microprotéines

La découverte récente d'ARN atypiques bouleverse la génétique. Ces longs ARN pourraient être plus nombreux que les gènes classiques. Alors qu'on les croyait non-codants, certains peuvent produire des microprotéines capables de contrôler l'expression du génome pour le développement de l'embryon.



L'ARN pri produit des microprotéines qui contrôlent le développement embryonnaire des insectes, notamment la formation des trichomes.

Les avancées du séquençage à très haut débit révèlent l'existence d'un nombre croissant de longs ARN dits non codant. On estime que l'homme en porterait entre 10 000 et 100 000. Contrairement à la plupart des gènes, ces ARNs ne produisent pas de protéines. Certains longs ARN non codants agissent sur l'expression des gènes, en modulant par exemple l'organisation de la chromatine, la transcription, la stabilité des ARN messagers ou encore leur traduction. En étudiant le développement des insectes, les chercheurs ont montré qu'un long ARN non codant produit en réalité des microprotéines, indispensables à différents programmes de différenciation au cours de l'embryogenèse.

Résultats surprenants

Tout ARN de grande taille (>1000 nucléotides) contient statistiquement de petits cadres de lecture (régions potentiellement traduites en protéines) qui pourraient produire de petits peptides de quelques acides aminés. On consi-

dère cependant qu'en dessous de 100 acides aminés, ces peptides ne sont pas fonctionnels. Ils sont systématiquement ignorés par l'annotation des génomes. En collaboration avec l'équipe de Yuji Kageyama (Okasaki, Japon), les travaux des chercheurs du CDB ont montré chez la drosophile que l'ARN considéré non codant *polished-rice (pri)* n'obéit pas à cette règle. En effet, cet ARN de 1500 nucléotides possède 4 petits cadres de lecture produisant des microprotéines de 11 à 32 acides aminés. La mise au point d'un système reconstitué de cellules en culture a permis de disséquer le mode d'action de l'ARN pri. Quatre mutations ponctuelles qui empêchent la production des microprotéines Pri, sans altérer le reste de l'ARN, sont suffisantes pour en inactiver la fonction. Réciproquement, un ARN synthétique capable de produire une microprotéine Pri peut remplacer l'ARN endogène pri. Ensemble, ces résultats surprenants démontrent qu'un ARN apparemment non codant fonctionne de fait via la synthèse de microprotéines.



François Payre et Serge Plaza, directeurs de recherche CNRS au Centre de biologie du développement (CBD, unité mixte UPS/CNRS)



Contact

serge.plaza@univ-tlse3.fr
francois.payre@univ-tlse3.fr

Contrôle du développement embryonnaire

Les embryons dépourvus des microprotéines pri, présentent de graves défauts, notamment du système respiratoire et de l'épiderme. L'équipe a décrypté comment ces microprotéines contrôlent le développement de l'épiderme, pour la formation d'extensions nécessaires à la locomotion (les trichomes). Ils ont identifié les bases moléculaires de la formation des trichomes, dont le rôle clé d'un facteur de transcription, Shavenbaby, qui contrôle l'expression de l'ensemble des gènes de la différenciation des trichomes. En couplant analyses *in vivo* et biochimie moléculaire, ils ont montré que les microprotéines Pri sont indispensables à l'activité de Shavenbaby. Ce dernier est initialement produit en une grande protéine répresseur de transcription. L'expression des microprotéines Pri déclenche le clivage de Shavenbaby, qui devient un activateur de transcription. Les microprotéines Pri, en régulant l'activité d'un facteur de transcription, assurent ainsi le contrôle temporel du programme de différenciation de l'épiderme. Ces découvertes révèlent l'importance insoupçonnée des microprotéines. Elles suscitent la recherche d'autres molécules de ce type requises pour le développement normal, et leur implication dans certaines maladies génétiques et virales. ■

Ces travaux ont bénéficié en particulier du soutien de l'ARC (Association pour la recherche sur le cancer).

Les microARNs bouleversent l'oncologie?

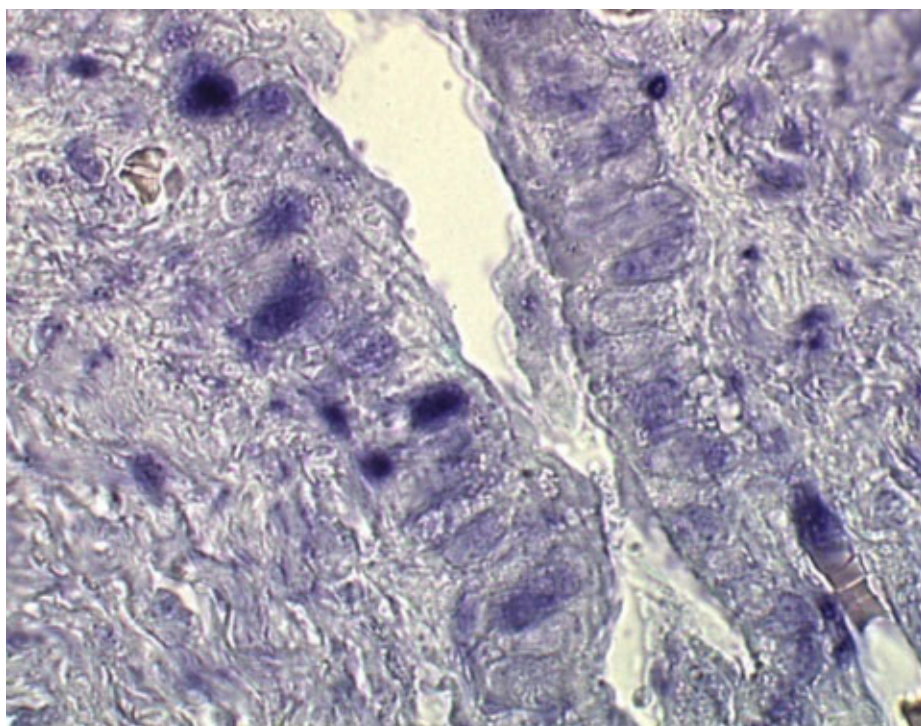
Biomarqueurs des cellules cancéreuses, les microARNs (miARNs) pourraient permettre de poser des diagnostics précoces. Certains se comportent aussi comme de véritables acteurs de l'oncogénèse, et pourraient devenir de nouvelles cibles thérapeutiques.

Au cours des 8 dernières années, 16 000 travaux de recherche sur les miARNs ont été publiés: les miARNs sont impliqués dans tous les aspects de la biologie tumorale. En oncologie, l'expression des miARNs permet de distinguer le tissu normal du tissu cancéreux et, de manière remarquable, de prendre le relais de l'analyse histologique afin de discriminer différents sous-types de cancer de nature inconnue. La grande stabilité des miARN dans les tissus et les fluides biologiques représente un avantage majeur, par rapport notamment aux ARN messagers (ARNm), ce qui augmente leur potentiel en tant que biomarqueurs surtout que leur abondance facilite leur détection. Il est également désormais bien établi qu'un seul miARN est capable de cibler très efficacement de multiples ARNm. Ainsi, on parle désormais de deux grandes classes de miARNs en oncologie: les oncomiRs, qui sont surexprimés dans les tumeurs et qui favorisent la progression

cancéreuse, opposés aux miARNs suppresseurs de tumeurs (ou tsmiRs) dont l'expression est perdue dans les tumeurs et qui jouent le rôle de régulateurs négatifs de la prolifération cellulaire. Ainsi, « la manipulation » d'un seul miARN permettrait la modulation simultanée de plusieurs signaux de transduction, en comparaison avec les approches classiques ne pouvant cibler qu'un seul gène ou une seule protéine. Ceci est particulièrement important pour éviter la clonalité des cellules tumorales.

Médecine personnalisée

Les miARNs apparaissent de plus en plus comme des biomarqueurs prometteurs pour une médecine personnalisée. Un exemple remarquable est la mise au point d'un test simple basé sur la détection de miR-210, qui a été récemment décrit comme pouvant prédire la réponse au tamoxifène chez les patientes atteintes de cancer du sein avec une précision comparable au test de



Mise en évidence de l'expression d'un oncomiR (coloration bleue foncée), miARN favorisant le cancer du pancréas.



Alix Vignolle Vidoni, technicienne au CHU, **Louis Buscail**, professeur UPS, **Hubert Lulka**, technicien UPS, **Coralie Cazals**, stagiaire IUT, **Pierre Cordelier**, directeur de recherche Inserm, **Barbara Bournet**, doctorante, **Naima Hanoun**, ingénieure d'études Inserm, **Jérôme Torrisani**, chargé de recherche Inserm, **Thoria Diab**, doctorante, **Lucrèce Dagnon**, stagiaire, tous au Centre de recherches de cancérologie de Toulouse (CRCT, unité mixte Inserm/UPS)



Contact

pierre.cordelier@inserm.fr

référence (OncotypeDX). D'un point de vue thérapeutique, le premier médicament ciblant un miARN est entré en phase 2 clinique pour le traitement de patients infectés par l'hépatite C. Il est donc tentant de spéculer que les miARNs seront utilisés en routine dans les essais cliniques, une fois que les résultats prometteurs de phases précoces auront été vérifiés lors d'essais cliniques à grande échelle.

Diagnostic précoce

À ce jour, le diagnostic précoce du cancer du pancréas est impossible. Par conséquent, 80 % des patients présentent une maladie avancée, et la médiane de survie globale est de 4 à 6 mois. Nous avons récemment démontré que l'expression des miARNs est fortement altérée à des stades précoces de la carcinogénèse pancréatique, que ce soit dans des modèles animaux transgéniques ou après analyse des pièces tumorales issues de patients. Cette altération est maintenue dans les phases avancées de la maladie. Nous recherchons désormais activement la présence de miARNs candidats dans les fluides biologiques comme autant de biomarqueurs de diagnostic spécifique et non invasif. Nous explorons aussi l'intérêt du dosage des miARNs dans le sang circulant comme marqueurs prédictifs de réponse et de suivi chez les patients traités par thérapie génique. Enfin, nous avons démontré que certains oncomiRs sont des cibles thérapeutiques qui permettent la chimiosensibilisation et par conséquent l'éradication de tumeurs pancréatiques expérimentales. Ces derniers travaux permettent de considérer, dans un futur proche, que les miARNs deviennent des acteurs incontournables en oncologie pancréatique. ■