

De beaux lendemains pour les HYER ?

Jean-Christophe Pagès

Unité UMR1301 RESTORE, Biologie cellulaire
CHU, Toulouse, France.

jean-christophe.pages@univ-tlse3.fr

Voilà des milliards d'années que les ARN nous façonnent, et quelques dizaines qu'ils nous fascinent. Le monde des ARN, à l'origine des génomes des êtres vivants contemporains, s'est structuré autour des propriétés physico-chimiques de ces molécules, en interaction avec les lipides et les acides aminés entre autres [1, 2]. La catalyse enzymatique, propriété centrale des ARN catalytiques ou « ribozymes », et l'identification du rôle de ces ribozymes dans la maturation des introns¹, ont ouvert la voie à la compréhension de l'émergence et de l'évolution des génomes des organismes. Les recherches de caractérisation de l'immense diversité des ARN se poursuivent [2].

La capacité des ribozymes à hydrolyser les liaisons des acides nucléiques concerne certains rétrotransposons² comme des introns du groupe II (GII). Les GII, composés de trois classes (A, B et C) [3], sont principalement présents chez les bactéries (génomes ou séquences plasmidiques), mais aussi dans les organelles d'origine endosymbiotique, mitochondries et chloroplastes chez les plantes et les champignons [2]. Sur le plan moléculaire, ces éléments

fonctionnent en quatre étapes : la transcription, l'excision, l'intégration et la transcription inverse (Figures 1A, 1B). Comme la majorité des ARN, ils possèdent une structure intramoléculaire conditionnée par leur séquence. Les GII se replient en tiges et boucles, leur conférant une structure de repliement en trois dimensions conservée. Les GII comportent six domaines (DI à DVI) (Figure 1C). Le domaine I permet la reconnaissance d'une séquence d'intégration dans l'ADN cible (*exon binding site*, EBS), les domaines II et III contribuent à la stabilité de l'ARN, le domaine IV contient un cadre ouvert de lecture (*open reading frame*, ORF) codant une endonucléase et une transcriptase inverse formant la maturase, les domaines V et VI permettent la formation d'un intermédiaire circulaire par transestérification (Lariat). Cette organisation est retrouvée dans les trois classes de GII, en notant néanmoins qu'il existe des GII dont la boucle IV est dépourvue d'ORF. Ces ORF-less GII étaient jusqu'ici considérés comme des formes dégénérées, incapables de mobilité génomique.

Le séquençage de l'ADN à haut débit combiné à l'analyse bio-informatique des banques de séquences, a permis au groupe de Jun-Jie Gogo Liu (*Beijing advanced innovation center for structural biology*, Pékin, Chine) [4] d'effectuer une recherche systématique des éléments transposables de type GII. Cette équipe s'est intéressée à la sous-classe des ORF-less GII C mobiles, qui n'avait pas encore été décrite. Liu et ses collaborateurs ont établi des critères de sélection incluant l'existence des domaines des

GII connus, une taille inférieure à 4 000 nucléotides, un domaine IV de moins de 200 nucléotides, et une présence sous forme de groupes (*cluster*) au sein de génomes, ce critère étant une signature d'activité. Dans les banques accessibles, ils ont identifié 469 candidats parmi 5 635 possibles, et en affinant le modèle, 93 séquences ont été retenues. Cette identification était un premier résultat probant et la confirmation de l'existence probable d'ORF-less GII C actifs. Ce critère impliquait alors un mode de répliation propre à ces éléments, indépendant d'une activité endonucléase de la maturase. La suite des travaux a donc visé à identifier les GII qui se propageaient au travers d'un clivage catalysé par une activité de ribozyme. La sélection de GII C ORF-less actifs en a alors ciblé 9 dotés d'un domaine I avec EBS et thermodynamiquement stables.

Pour confirmer la propagation indépendante d'une maturase, les ARN devaient pouvoir cliver un ADN. Les auteurs ont mis en œuvre une approche séquentielle de validation des éléments identifiés par l'analyse informatique. Ils ont commencé par tester la coupure d'un ADN simple brin (ADNsb). Cette activité a été constatée pour six des neuf candidats. La coupure étant spécifique d'un EBS d'intron, cela suggérait aussi qu'un fragment de la séquence ARN du GII guidait la reconnaissance de la cible ADN. Une analyse plus fine indiquait que la coupure était située en 3' du site d'hybridation de l'ARN à l'ADN, signant la présence d'une séquence de reconnaissance (*target recognition site*, TRS) (Figure 1C). Les études se sont poursuivies sur les deux séquences les plus efficaces, nommées

¹ Les introns du groupe I, identifiés en 1982 dans les ARN de génomes bactériens, phagiques, viraux, d'organelles, et dans les ARN ribosomiques nucléaires des champignons, des plantes et des algues, ont une activité d'excision auto-catalysée, en l'absence de protéine.

² Les rétrotransposons sont des éléments intragénomiques mobiles se propageant par transcription inverse d'un intermédiaire ARN qui s'intègre dans le génome. Ils sont une source importante de diversité génétique.



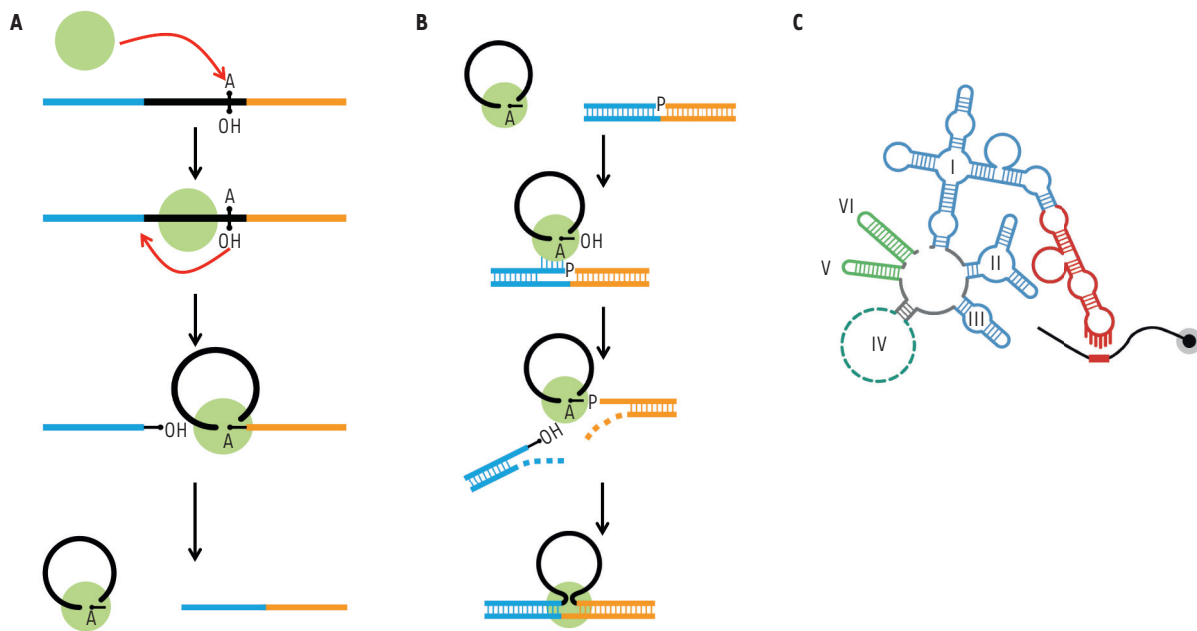


Figure 1. Mécanismes de mobilité et structure en domaines des éléments transposables GII. **A.** La maturase permet l'hydrolyse de l'intron, formant le Lariat, et la liaison des deux exons de l'ARN porteur du GII. **B.** Le complexe Lariat/maturase s'intègre et forme une jonction exon-intron. La maturase hydrolyse les brins d'ADN et permet la liaison du Lariat suivie de la transcription inverse (non représentée). **C.** Structure des GII. Les domaines (D) I à VI sont notés en chiffres romains. L'interaction avec l'intron résulte d'un appariement d'une boucle en épingle à cheveux de DI avec l'intron (représenté par un trait noir, avec un rectangle rouge indiquant sa région d'appariement) par un TRS (*target recognition site*) (illustrations : Nathan Pagès d'après [3] et [4]).

HYER (*hydrolytic endonucleolytic ribozyme*). Pour ces deux séquences (HYER1 et HYER2), la capacité à cliver de l'ADN double brin a été validée *in vitro* en utilisant des plasmides (Figure 2A). Toutefois, en l'absence d'activité hélicase, seules les régions simple brin (*bubble*) de ces ADN étaient clivées. En utilisant un élégant système de contre sélection dans un modèle bactérien, l'activité d'hydrolyse a également été montrée *in vivo* après sélection. Dans une cellule eucaryote, seule la séquence HYER1 est active dans un système de sélection reposant sur un recalage de cadre de lecture au sein d'un gène de résistance à la puromycine : en l'absence de coupure, il n'y a pas de synthèse d'une puromycine N-acétyltransférase fonctionnelle. Par séquençage et utilisation de variants inactifs de HYER1, les auteurs ont montré que la coupure est bien spécifique de la séquence TRS. Sur le plan moléculaire, ils ont identifié des inser-

tions/délétions (INDEL) allant de 1 à 85 nucléotides, avec une faible efficacité de coupure (0,09 à 0,2 %). Les auteurs ont ensuite réalisé une analyse de la structure moléculaire des formes actives de HYER1. HYER1 apparaît former des homodimères de symétrie plane, le plan de symétrie résultant de l'hybridation de palindromes des séquences nucléotidiques des domaines I et III (Figure 2B). L'analyse a montré également la liaison de l'élément à l'ADN, et a confirmé la présence d'un cation divalent (Mg^{2+}) au site de coupure. Ces données structurales montraient que les TRS n'étaient pas engagés dans la structuration en trois dimensions de l'ARN des HYER, un résultat important puisqu'il suggérait que l'on pouvait modifier ces TRS sans altérer la structure ou la fonction du HYER. Les auteurs ont alors entrepris une série de substitutions de nucléotides et de tests de séquence

et de taille des TRS, afin de voir si l'on pouvait cibler la coupure à une séquence d'ADN choisie. Dans une bactérie, un TRS de 10 nucléotides est fonctionnel et spécifique, et la présence d'un seul mésappariement avec la séquence cible inhibe le clivage. *In vitro*, les auteurs ont pu éprouver la souplesse du système en étendant la taille de reconnaissance du TRS : la capacité de coupure est alors maintenue, et spécifique car tout mésappariement inhibe la coupure. Avec cet ensemble de données, la possibilité de créer un nouvel outil de modification des génomes prend forme. Néanmoins, ces coupures ne concernent qu'une unique séquence qui ne devait pas être engagée dans un appariement ADN/ADN. Cela explique la faible efficacité de sélection constatée *in vivo* sur les cellules humaines avec le système de tolérance à un antibiotique. Du fait de la structure de HYER1 et de la formation des homodimères par appariement

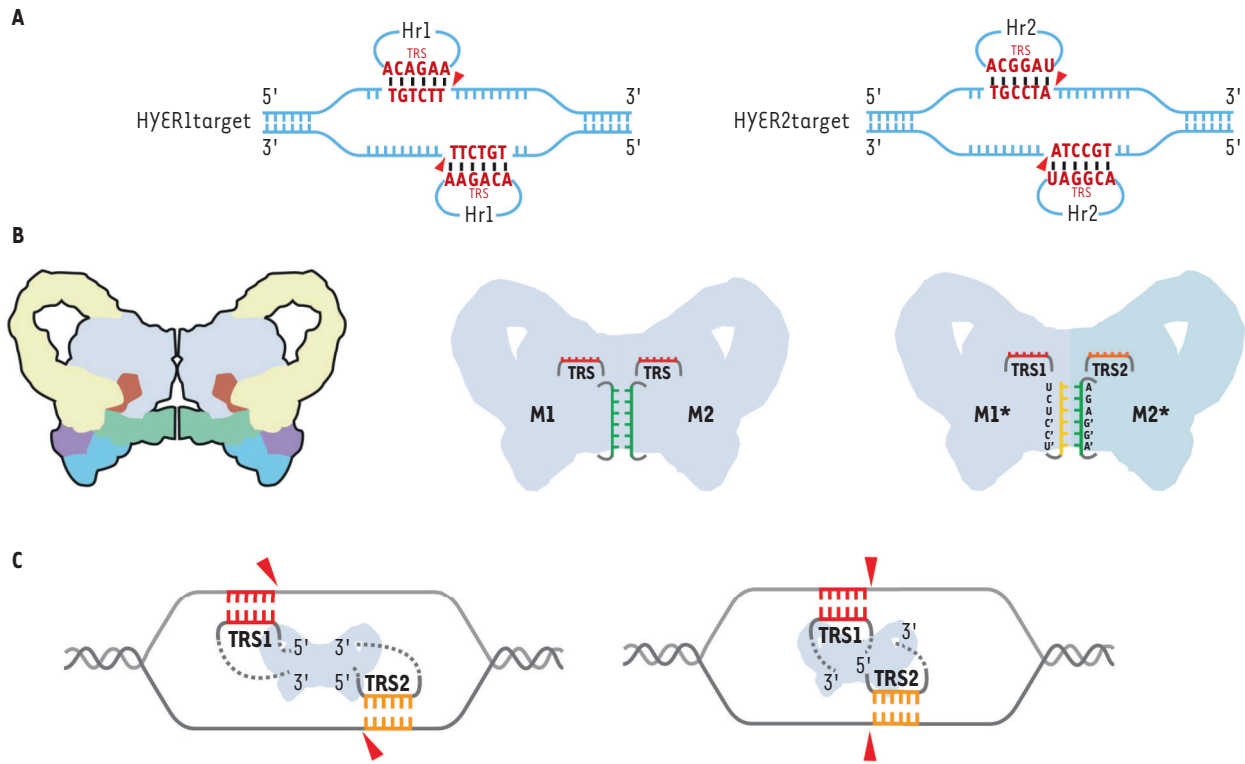


Figure 2. Coupure ciblée de l'ADN par les HYER. **A.** Séquence des deux HYER les plus actifs : après avoir validé la coupure d'un ADN synthétique simple brin, les auteurs testent les coupures d'ADN double brin *in vivo* dans les bactéries et les cellules humaines. **B.** Structure dimérique des HYER. Chaque domaine, DI à DVI, est identifié par une couleur. Les DI (en gris) et DIII (en vert) contiennent les zones d'interaction entre les deux ARN du GII. Visualisation des TRS (*target recognition site*) et de la séquence d'appariement des homodimères M1 et M2. Forme hétérodimérique M1 et M2*. **C.** Un hétérodimère doté de deux TRS différents coupe un ADN double brin et produit des extrémités 5' ou 3' (non montrées) sortantes, ou des extrémités franches (illustrations : Nathan Pagès d'après [4]).

d'un palindrome interne, les auteurs ont testé la modification des zones d'hybridation afin d'obtenir des hétérodimères (Figure 2B) : l'intérêt était de pouvoir disposer de deux séquences liant deux TRS différents, ce qui ouvrait à la possibilité de couper les deux brins complémentaires d'un ADN portant des séquences différentes sur chaque brin. Le changement du palindrome interne a permis la formation d'hétérodimères et, en jouant sur la flexibilité d'orientation des TRS, d'obtenir une coupure sur les deux brins d'un ADN (Figure 2B). Selon la séquence de liaison du TRS, il est possible d'obtenir des coupures franches, ou des coupures 3' ou 5' sortantes (Figure 2C). Ainsi, la formation d'un nouveau type d'endonucléase se concrétisait. Notons que cette dernière

série d'expériences n'a été réalisée qu'*in vitro*, avec un ADN synthétique. Ce travail de recherche, remarquable par son approche et son ampleur, ouvre la voie à la mise à disposition d'un nouvel outil de coupure des génomes, dont l'intérêt majeur est la simplicité moléculaire car il ne repose que sur une ou deux molécules d'ARN ! Qui plus est, avec l'avantage d'une taille réduite (0,6 kb), ce qui permet d'envisager des modes de synthèse et d'administration aisés et peu coûteux. Mais avant d'envisager une utilisation plus large des HYER, il reste à confirmer leur activité sur un grand nombre de cibles et à tester l'efficacité de la coupure double brin dans des systèmes *in vivo*. Pour ces deux points, l'ouverture de l'ADN cible est déterminante, ce qui nécessite

une activité hélicase que les HYER ne possèdent actuellement pas. La possibilité de les en doter par évolution moléculaire est une hypothèse aussi belle qu'audacieuse [4], mais peut-être illusoire. Néanmoins, l'identification des ORF-less GII montrent que leur insertion est possible aux sites d'ouverture des brins d'ADN, lors de la réplication des génomes par exemple [5]. Il conviendra d'évaluer également le degré de spécificité de la coupure de l'ADN cible par les HYER, qui conditionne en grande partie leur utilisation. En effet, les coupures hors cible comportent un risque pour les applications médicales, et rendent nécessaire une vérification, avant leur éventuelle élimination, pour les applications en recherche et en agronomie. Quoi qu'il en soit, cette technique ne devrait

pas se substituer aux systèmes CRISPR/Cas de modification du génome, dont la composante protéique peut être couplée à une grande diversité d'activités enzymatiques, ce qui élargit le champ de leurs utilisations [6]. Cependant, dans une perspective de combinaison de techniques, rien n'interdit d'imaginer la possibilité d'associer les techniques HYER et CRISPR/Cas.

Cette diversité moléculaire des éléments mobiles du génome étend, avec les HYER, à de nombreux ARN les précurseurs qui auraient contribué à l'émergence du vivant. Centrant leur conclu-

sion sur l'évolution des organismes dans le monde de l'ARN, les auteurs suggèrent des pistes expliquant la sélection et le maintien des HYER dans les génomes contemporains. ♦

GII-derived tools link HYER to higher organisms' genome editing

LIENS D'INTÉRÊT

L'auteur déclare n'avoir aucun lien d'intérêt concernant les données publiées dans cet article.

RÉFÉRENCES

1. Müller F, Escobar L, Xu F, et al. A prebiotically plausible scenario of an RNA-peptide world. *Nature* 2022 ; 605 : 279-84.
2. Fine JL, Pearlman RE. On the origin of life: an RNA-focused synthesis and narrative. *RNA* 2023 ; 29 : 1085-98.
3. McNeil BA, Semper C, Zimmerly S. Group II introns: versatile ribozymes and retroelements. *WIREs RNA* 2016 ; 7 : 341-55.
4. Liu ZX, Zhang S, Zhu HZ, et al. Hydrolytic endonucleolytic ribozyme (HYER) is programmable for sequence-specific DNA cleavage. *Science* 2024 ; 383 : eadh4859.
5. Waldern J, Schiraldi NJ, Belfort M, Novikova O. Bacterial group II introns genomic neighborhoods reflect survival strategies: hiding and hijacking. *Mol Biol Evol* 2020 ; 37 : 1942-8.
6. Schmidt TZN, Berarducci B, Konstantinidou S, Raffa V. CRISPR/Cas9 in the era of nanomedicine and synthetic biology. *Drug discovery today* 2023 ; 28 : 103375.