

Rapport d'activité (janvier 2015 – décembre 2019)

Hervé Seitz

Table des matières

1	<i>Curriculum vitæ</i>	3
1.1	Informations générales	3
1.2	Études et expérience post-doctorale	3
1.3	Enseignement	3
1.4	Activités d'expertise	4
1.5	Publications dans des revues à comité de lecture	5
1.6	Principales communications orales depuis la création de mon équipe	5
1.7	Vulgarisation	6
1.8	Prix et distinctions	7
1.9	Activités non professionnelles : ultra-marathon	7
2	Recherche scientifique	8
2.1	La fonction biologique des microARN a été sur-évaluée	8
2.2	Cibles fonctionnelles du miARN <i>bantam</i> chez la Drosophile	10
2.3	Contrôle de la prolifération cellulaire par miR-34	14
2.4	Évolution de la machinerie de l'interférence à ARN chez les Animaux	17
2.5	Recherche d'inducteurs de la dégradation des microARN	20
2.6	Déterminants génomiques et épigénétiques des origines de réplication au cours du développement embryonnaire de <i>Cænorhabditis elegans</i>	22
2.7	Collaborations	22
2.8	Place de ma recherche dans mon unité	23
2.9	Mobilités	23
2.10	Distinctions scientifiques	23
2.11	Publications scientifiques	23
2.11.1	Revue à comité de lecture	23
2.11.2	Conférences invitées dans des congrès	24
2.11.3	Actes de colloques à comité de lecture	24
2.11.4	Publications dans des revues sans comité	24
2.11.5	Communications à des congrès, symposium	24
2.11.6	Séminaires, <i>workshops</i>	24
2.11.7	Livres et ouvrages	25
2.11.8	Chapitres d'ouvrage	25
2.11.9	Logiciels	25
2.11.10	Autres	25

3	Enseignement, formation et diffusion de la culture scientifique	25
3.1	Thèses dirigées	25
3.1.1	Sophie Mockly	25
3.2	Post-doctorants encadrés	26
3.2.1	Natalia Pinzón	26
3.2.2	Blaise Li	26
3.3	Participations à l'enseignement	27
3.4	Participation à l'organisation de conférences	27
3.5	Participation à la diffusion de l'information scientifique et technique vers des publics non spécialistes	27
3.6	Interventions dans la presse écrite et audiovisuelle	28
3.7	Participation à des travaux d'expertise	28
4	Transfert technologique, relations industrielles et valorisation	29
4.1	Contrats de recherche	29
4.2	Activités de consultance	29
5	Encadrement, animation et management de la recherche	29
5.1	Direction d'équipe	29
5.2	Autres responsabilités ou activités collectives au sein du CNRS	30
6	Références	31

1 *Curriculum vitæ*

1.1 Informations générales

Né le 13 décembre 1977 à Angers (Maine-et-Loire); email : herve.seitz@igh.cnrs.fr.

Site web de l'équipe : <https://www.igh.cnrs.fr/fr/recherche/departements/27>

Adresse professionnelle : IGH (UMR 9002 CNRS et univ. de Montpellier), 141 rue de la Cardonille, 34396 Montpellier.

Profil ResearchGate : https://www.researchgate.net/profile/Herve_Seitz2

Identifiant ORCID : 0000-0001-8172-5393.

Scores bibliométriques (nombre de publications, h-index et choses de ce genre) : sur <https://publons.com/researcher/1288691/herve-seitz/>.

1.2 Études et expérience post-doctorale

- 1997 : admis aux trois Écoles normales supérieures, dont l'É.N.S. de la rue d'Ulm avec le rang : 2.
- 1997–2000 : Magistère de chimie, option chimie-biologie, de l'É.N.S. (1997-1999 : licence puis maîtrise de chimie ; 1999-2000 : D.E.A. « génétique moléculaire et cellulaire » des universités Paris VI, Paris XI et UVSQ).
- 2001–2004 : thèse (financement « allocation couplée »); titre : « Empreinte génomique parentale et petits ARN non-codants » (université Toulouse III Paul Sabatier) (directeur de thèse : Jérôme Cavallé; soutenance le 25 octobre 2004). Manuscrit disponible sur : <http://www.normalesup.org/~seitz/These.pdf>
- Janvier 2005–janvier 2009 : stage post-doctoral à University of Massachusetts Medical School (Worcester, Massachusetts, États-Unis); directeur des travaux : Phillip D. Zamore.
- 2008 : classé 1^{er} au concours de chargé de recherche de deuxième classe de la section 21 du CNRS; poste pris en janvier 2009.
- 3 novembre 2009 : Habilitation à diriger les recherches. Manuscrit disponible sur : <http://www.normalesup.org/~seitz/HDR.pdf>
- 2010 : lauréat du Career Development Award (CDA n°00017-2010-C) de Human Frontier Science Program.
- 2010 : classé 1^{er} pour le recrutement d'un jeune chef d'équipe à l'IGH (UPR 1142 du CNRS; Montpellier); installation à l'IGH en octobre 2011.
- 2011 : lauréat du programme ATIP-Avenir (contrat initié en mars 2012).
- 2013 : promu chargé de recherche de 1^{ère} classe.
- 2017 : promu « chef d'équipe senior », permanent, à l'IGH, après évaluation par un jury international.
- 2019 : classé 2^{ème} au concours de directeur de recherche de deuxième classe de la section 21 du CNRS.

1.3 Enseignement

Cours magistraux sur les petits ARN régulateurs eucaryotiques (d'une à trois interventions par an, depuis 2007, pour étudiants en M2R ou en thèse, aux universités de Toulouse, Paris, Strasbourg, Lyon, Rennes et Montpellier; dans deux universités étrangères en 2014 : *National Taiwan university* et *Yang Ming university*; à l'École normale supérieure : cours de 6h en L3 chaque année de 2013 à 2018, cours de 2h en M1 en 2016, 2017 et 2019).

Intervention d'une heure au *International course on Non-coding genome* de l'Institut Curie, le 11 décembre 2012, puis à nouveau le 8 février 2018; intervention d'1h30 à la journée *DE[^]CODAGE* de l'INRA de Castanet-Tolosan, le 26 avril 2013; intervention de 45 min au symposium *Statistical*

analysis of RNA-seq data : advances and challenges de la plate-forme de transcriptomique de l'Institut Pasteur, le 26 novembre 2013.

Je pense que la pratique de notre métier devrait être coopérative; c'est pourquoi je partage systématiquement les programmes que j'écris avec les collaborateurs qui ont sollicité mon aide, et je leur apprend à s'en servir. Je propose des formations formelles ou informelles en bio-informatique et statistiques, des méthodes très utiles pour les biologistes expérimentalistes. Je mets ces formations en accès libre pour qu'elles servent au plus grand nombre : la partie théorique du cours de statistiques a été publiée dans les numéros de juillet et d'octobre 2010 de *Regard sur la Biochimie*, et les diaporamas de toutes ces formations sont accessibles sur <http://www.igh.cnrs.fr/equip/Seitz/equipe-tutoriels.html>.

Membre du jury de recrutement d'un maître de conférence à l'université Toulouse III Paul Sabatier (poste 64MCF0300, en 2013). Membre d'un jury de thèse de pharmacie et de 23 jurys de thèse scientifique (19 fois en tant que rapporteur, 4 fois en tant qu'examinateur), pour des étudiants d'institutions françaises et étrangères (*Cold Spring Harbor Laboratory, Guangzhou university, Caltech, university of Queensland, ...*). Rapporteur dans deux jurys d'HDR (S. Chambeyron, 2011 ; I. Duroux-Richard, 2018).

1.4 Activités d'expertise

Rapporteur pour des journaux dans le domaine de l'ARN, de la bio-informatique et de la génétique moléculaire (*Current Biology, EMBO Reports, EMBO Journal, Genome Biology, Genome Research, Nucleic Acids Research, RNA Biology, ...*) (la plupart de mes activités de *reviewing* pour des journaux depuis 2015 sont visibles sur <https://publons.com/researcher/1288691/herve-seitz/>). Éditeur invité pour un numéro spécial de *Frontiers in Genetics* sur les microARN.

Membre du *grant panel* « *Governance Leadership Team* » du *Cefic Long-range research initiative* pour les domaines « omics, epigenetics and non-coding RNAs » pour l'amélioration des procédures d'évaluation des risques toxicologiques (2016). *Chairman* du comité de sélection pour l'appel d'offre "LRI C6 - Omics and Read across" du Cefic LRI (2017).

Membre du comité de programme du congrès JOBIM (Journées Ouvertes en Biologie, Informatique et Mathématiques) 2012 et 2013. Membre du comité de sélection des projets de la plateforme ABIC de Montpellier (2016–2017).

Rapporteur pour des agences de financement nationales et internationales :

- KU Leuven Research Council (2019),
- ERC Starting grant (2017),
- Swiss national science program (2017 et 2018) ;
- Cefic Long range initiative (2017) ;
- ANSES (2017),
- Human Frontier Science Program (2016),
- l'ANR (2012 : panel SVSE2 ; 2014 : panel CE10 ; 2016 : panel CE12),
- l'IUF (2015),
- l'Idex Sorbonne universités (2014),
- la FRM Alsace (2011),
- le GIS IBiSA (2010).

Activité de consulting pour Medesis Pharma S.A. (compagnie spécialisée dans le développement de vecteurs pour l'adressage *in vivo* de composés actifs ; je les ai aidés à choisir des siRNA contre des gènes d'intérêt pour répression *in vivo* chez la Souris) : de mars à mai 2012, puis de mars à mai 2019.

Porteur de l'école thématique *InteRNAt* du CNRS sur les petits ARN régulateurs eucaryotiques. Cette école thématique s'est tenue du 6 au 10 octobre 2019, et le comité d'organisation rassemblait, outre moi-même, S. Chambeyron (Montpellier), K. Mochizuki (Montpellier), A. Verdel (Grenoble), J. Cavallé (Toulouse), M. van den Beek (Paris) et S. Pfeffer (Strasbourg).

- 3 mars 2016 à l'*ECETOC workshop "ncRNAs and risk assessment science"* à Málaga (Espagne) (**conférencier invité**).
- 12 février 2016 au *SMPGD (statistical methods for post-genomic data) annual workshop* à Lille (**keynote speaker**).
- 4 juin 2015 à l'*ESPE science school "Non-coding RNAs in paediatric endocrinology"* à Annecy (**conférencier invité**).
- 9 juin 2014 au *Taiwan-France mini-symposium on regulatory RNAs* à Taipei (Taïwan) (**conférencier invité**).
- 2 février 2014 au *Keystone Symposium "RNA Silencing (A9)"* à Seattle (États-Unis).
- 26 novembre 2013 à l'atelier *Statistical analysis of RNA-seq data : advances and challenges* de l'institut Pasteur à Paris (**conférencier invité**).
- 22 mars 2013 au *Keystone Symposium "RNA Silencing (C7)"* à Whistler (Canada).
- 11 décembre 2012 à l'*International course on non-coding genome* de l'institut Curie à Paris (**conférencier invité**).
- 13 octobre 2012 au *BTCure miRNA workshop* à Montpellier (**conférencier invité**).
- 12 septembre 2012 à la *Annual French conference on Drosophila* à Clermont-Ferrand (**conférencier invité**).
- 24 avril 2012 à la *Schlumberger Foundation conference "Small Silencing RNA Biology and Mechanism"* à Tourtour (**conférencier invité**).

1.7 Vulgarisation

Articles de vulgarisation : *Pour la Science* (1 article en 2005, 1 article en 2013). Interventions dans les médias : *Néon* (1 contribution en 2014), *Le Monde* (1 intervention en 2015), <http://www.atlantico.fr> (1 contribution en 2015, 4 contributions en 2016, 2 contributions en 2018, 1 contribution en 2019).

Activités réalisées pour l'ANSES : Correction d'une note publiée par le Bulletin de veille scientifique n°30 (octobre 2016), rédaction de quatre notes pour le Bulletin de veille scientifique (n°33, 34, 35 et 36). Évaluation d'un projet de recherche pour l'appel d'offre PNR EST de l'ANSES (2017).

Conférences pluridisciplinaires : présentation de la biotechnologie CRISPR/Cas9, de ses implications pratiques et éthiques, au colloque *Quelles limites pour les technosciences en santé ?* de l'école de droit de l'université Clermont Auvergne, le 13 mars 2018 à Clermont-Ferrand (actes publiés dans le n°15 de la Revue du Centre Michel de l'Hospital, en septembre 2018). Présentation du développement du séquençage personnalisé, au colloque *Santé et droit de choisir* de l'école de droit de l'université Clermont Auvergne, le 28 novembre 2019.

Conférences publiques : Le grand public est très intéressé par la génétique moléculaire. Alors que la génération du *baby-boom* n'a pas entendu parler d'ADN ou de chromosomes au lycée, les électeurs sont amenés à se prononcer sur les OGM, le clonage, à faire des dons pour la thérapie génique, ... Le public est demandeur d'informations sur ce qui se fait dans les laboratoires, et il est de notre devoir (en tant que chercheurs publics) de partager avec les citoyens le résultat des travaux menés en leur nom. J'essaie donc de contribuer à la diffusion de l'information scientifique vers le grand public. Je donne régulièrement des conférences (« Qu'est-ce que la génétique ? », « Comment l'ADN détermine le vivant », « Sur la notion de race. Le point de vue du biologiste. », « Utilisation de la génétique de la Drosophile pour explorer les mécanismes biologiques ») pour le public de la Fête de la science, devant des groupes de scolaires, pour l'Université du Tiers-Temps à Montpellier (conférences pour le grand public, essentiellement retraité), pour le Café des sciences d'Avignon, pour l'association « Rando loisirs culture » de Grabels, et pour le Visiatome de Marcoule. Les diaporamas de ces présentations sont accessibles sur le site web de mon équipe, et j'ai distribué le lien aux auditeurs pendant la présentation.

1.8 Prix et distinctions

- Prime d'encadrement doctoral et de recherche du CNRS (novembre 2019).
- Prix Maurice Nicloux de la SFBBM (2018).
- Prime d'excellence scientifique du CNRS (décembre 2010, puis novembre 2014).
- Bourse post-doctorale « Long-term fellowship » de HFSP (LT005752005-L).
- Prix d'Académie de l'Académie des sciences, inscriptions et belles-lettres de Toulouse (2005).
- Bourse post-doctorale « Long-term fellowship » de l'EMBO (ALTF 910-2004).

1.9 Activités non professionnelles : ultra-marathon

Année	Résultats
2019	Vainqueur des 100 km de Millau (champion d'Occitanie) (7ème meilleur performer français)
2018	Médaille de bronze au championnat de France de 100 km, vainqueur des 100 km de Millau (champion d'Occitanie) (3ème meilleur performer français)
2017	3ème des <i>100 km del Passatore</i> , 4ème des 100 km de Millau (vice-champion d'Occitanie) (7ème meilleur performer français)
2016	Vainqueur des 100 km de Belvès, vainqueur des 100 km de Millau (15ème meilleur performer français)
2015	3ème des 100 km de Belvès, vainqueur des 100 km de Millau (11ème meilleur performer français)
2014	2ème des 100 km de Millau (16ème meilleur performer français)
2013	2ème des 100 km de Millau (16ème meilleur performer français)
2012	Médaille de bronze au championnat de France de 100 km, 3ème des 100 km de Millau (12ème meilleur performer français)
2011	5ème des 100 km de Millau (22ème meilleur performer français)
2010	10ème des 100 km de Millau (52ème meilleur performer français)

2 Recherche scientifique

L'essentiel de nos travaux ont porté sur l'évaluation de l'importance fonctionnelle des interactions entre microARN (« miARN ») et cibles : en mêlant biologie moléculaire et bio-informatique, nous avons montré que les méthodes actuellement utilisées pour l'identification de cibles régulatrices de miARN sont lourdement contaminées par des faux positifs (section 2.1). Nous avons ensuite initié deux nouveaux projets, qui découlent de ces considérations : sur deux systèmes biologiques distincts, nous souhaitons mesurer la contribution d'interactions miARN/cibles individuelles au phénotype globalement contrôlé par le miARN (sections 2.2 et 2.3).

Le reste de nos travaux concerne également la fonction biologique de voies de régulation par les petits ARN, et la fonctionnalité non-codante du génome : nous avons montré la grande flexibilité évolutive de la machinerie d'interférence à ARN chez les Animaux (section 2.4), et (en collaboration avec le laboratoire de Marcel Méchali), identifié les déterminants génomiques et épigénétiques des origines de réplication au cours du développement embryonnaire de *Cænorhabditis elegans* (section 2.6). Notre projet le plus récent concerne la régulation des miARN eux-mêmes par des ARN-cibles particuliers : nous avons identifié bio-informatiquement des ARN qui semblent pouvoir contrôler l'abondance de certains miARN dans les neurones corticaux de Souris, et nous sommes en train de le vérifier expérimentalement (section 2.5).

Enfin, nous sommes fréquemment sollicités pour des collaborations, par des collègues qui recherchent de l'aide sur la biologie des petits ARN (section 2.7).

2.1 La fonction biologique des microARN a été sur-évaluée

Le contenu de cette section a été publié [Pinzón et al., 2017] : le lecteur est renvoyé à cet article pour les détails, et nous ne donnerons ici qu'un résumé synthétique.

Les miARN sont de petits ARN régulateurs qui reconnaissent des ARN-cibles spécifiques par complémentarité de séquence : ces cibles sont ensuite réprimées par les protéines effectrices associées au miARN [Iwakawa et Tomari, 2015]. La complémentarité entre miARN et cible n'a pas besoin d'être parfaite pour être fonctionnelle : des mésappariements sont bien tolérés, surtout sur l'extrémité 3' du miARN ; le segment délimité par les nucléotides 2 à 7 (comptés à partir du 5'), appelé « graine », semble quant à lui être le déterminant principal de reconnaissance des cibles : il a généralement besoin d'être parfaitement apparié à l'ARN-cible [Bartel, 2009].

Il existe actuellement deux grands types de méthodes pour identifier les cibles directes de miARN : la recherche d'interactants physiques du complexe effecteur des miARN (par exemple, un *cross-link* covalent, suivi d'une immuno-précipitation stringente et du séquençage à haut débit des ARN immuno-précipités ; [Mittal et Zavolan, 2014]) ; et la recherche bio-informatique de sites de complémentarité parfaite aux graines de miARN¹ phylogénétiquement conservés (partant du principe que si ces sites sont conservés, ils doivent être fonctionnels) [Friedman et Burge, 2014]. Si ces deux types de méthodes ne s'accordent pas sur l'identité exacte des cibles de miARN, elles tendent en revanche, toutes les deux, à proposer de nombreuses cibles : chaque miARN reconnaîtrait des centaines ou milliers d'ARNm différents, et, collectivement, les miARN réguleraient plus de 60% des gènes codants humains [Friedman *et al.*, 2009]. En revanche, l'effet répresseur des miARN sur chaque cible est très modeste (en général, quelques pourcents de répression, et jamais plus d'un facteur 2 ; [Selbach *et al.*, 2008, Baek *et al.*, 2008]).

Les approches bio-informatiques ont l'inconvénient de n'être que des prédictions, et les approches expérimentales à haut débit ont l'inconvénient de ne pas interroger la fonctionnalité biologique des interactions détectées : il est en effet important de se rappeler que la cellule est un environnement très concentré, où les macromolécules passent leur temps à interagir les unes avec les autres, au gré des collisions aléatoires, et ces interactions sont plus ou moins stables, en fonction de la présence de

1. Cette recherche est habituellement restreinte aux 3' UTR, où les miARN sont les plus efficaces [Bartel, 2009].

charges électriques sur les surfaces en contact. De très nombreuses interactions inter-moléculaires sont donc des événements inévitables, dans un environnement aussi concentré qu'une cellule, et elles ne jouent pas nécessairement de fonction biologique : si elles n'avaient pas lieu, l'organisme ne serait pas moins apte. Le moyen le plus simple, couramment utilisé pour vérifier la fonctionnalité biologique d'un événement moléculaire donné, consiste à mesurer sa conservation phylogénétique : si une interaction a été sélectivement préservée au cours de l'évolution, on considère qu'elle est biologiquement utile. C'est à ce titre que les programmes informatiques de prédiction de cibles sont très utilisés, même dans les études qui ont identifié expérimentalement des cibles de miARN : sans avoir à muter le site d'interaction, et à évaluer d'éventuels phénotypes biologiques après perte de l'interaction, la génomique comparative permet de s'assurer que l'interaction est biologiquement importante, et que ce n'est pas simplement une collision aléatoire sans fonction biologique.

Comme décrit dans mon précédent rapport d'activité à vague (janvier 2009 – décembre 2013), nous avons observé que, sur un système biologique particulier (le miARN miR-223 dans les neutrophiles de Souris), les cibles ainsi prédites informatiquement semblent fonctionnellement insensibles à l'action répressive de miR-223 : leur expression est très variable de souris à souris, dans une souche sauvage ne présentant pas de phénotype particulier. Pour environ 80% de ces gènes, cette variabilité inter-individus est d'ampleur plus large que l'effet répressif de miR-223, ce qui signifie que des fluctuations d'expression de ces gènes plus amples que l'effet du miARN semblent sans conséquence, et n'ont pas été éliminées par la sélection naturelle.

Ce phénomène ne semble pas spécifique à miR-223 ni aux neutrophiles de Souris : nous avons montré que chez l'Homme, les gènes connus pour être haplo-insuffisants [Dang *et al.*, 2008] portent des sites de complémentarité aux miARN plus conservés que les autres gènes, ce qui semble confirmer que les gènes dont la fonction biologique est la plus sensible à leur niveau d'expression, sont ceux pour lesquels la régulation par les miARN est la plus susceptible d'être biologiquement importante. De manière similaire : il est possible d'évaluer la probabilité, pour un gène humain, d'être haplo-insuffisant [Huang *et al.*, 2010], or les gènes pour lesquels cette probabilité est la plus forte sont aussi ceux dont les sites de reconnaissance de miARN sont les mieux conservés.

Nous avons donc confirmé le paradoxe que je présentais dans mon précédent rapport d'activité à vague (janvier 2009 – décembre 2013) : de nombreux gènes semblent fonctionnellement insensibles aux miARN (ils sont réprimés, mais si peu que cette répression ne provoque pas de phénotype macroscopique, ses conséquences sont annulées par les mécanismes d'homéostasie). Pourtant, toutes les cibles prédites de miARN portent des sites de fixation phylogénétiquement conservés, ce qui semble indiquer qu'ils sont fonctionnels ; pourquoi auraient-ils été conservés, si ces interactions ne contrôlent pas de phénotype physiologique, susceptibles d'être soumis à sélection naturelle ?

J'avais précédemment proposé [Seitz, 2009] que ces ARN pouvaient titrer les miARN : en s'y appariant, ils les empêcheraient d'aller réguler d'autres cibles (fonctionnellement sensibles à l'effet du miARN, elles), et cette fonction de modulation des miARN pourrait expliquer leur conservation phylogénétique.

Plusieurs autres laboratoires ont évalué cette possibilité, et il semble que les ARNm soient, dans l'immense majorité des cas, trop peu abondants pour pouvoir réellement affecter la disponibilité des miARN [Ebert et Sharp, 2010, Garcia *et al.*, 2011, Wee *et al.*, 2012, Denzler *et al.*, 2014, Broderick et Zamore, 2014, Bosson *et al.*, 2014, Denzler *et al.*, 2016]. Aucune de ces études n'avait cependant mesuré directement la concentration intracellulaire des miARN et des ARNm soupçonnés de les titrer. Nous avons réalisé cette expérience, dans des cellules de Souris en culture, les cellules C2C12. Il est possible d'induire la différenciation de ces cellules en fibres musculaires, et au cours de cette différenciation, les cellules commencent à exprimer les miARN spécifiques des fibres musculaires miR-1a/miR-206 et miR-133 ; la famille miR-1a/miR-206 est d'ailleurs impliquée dans le contrôle de la différenciation [Goljanek-Whysall *et al.*, 2012]. La différenciation des cellules C2C12 constitue donc un système bien adapté à la mesure d'une éventuelle compétition entre ARNm pour les miARN : à une date donnée dans la différenciation, c'est un échantillon relativement homogène de cellules qui se trouvent dans un état transcriptionnel similaire ; et les miARN semblent y remplir

une fonction biologique mesurable à l'échelle cellulaire.

Nos expériences montrent que quelques ARNm particulièrement abondants peuvent effectivement titrer ces miARN : si on sélectionne les ARNm dont la perte du site d'interaction au miARN devrait augmenter la concentration de miARN libre d'au moins 10% à l'un des trois stades de différenciation analysés, deux ARNm peuvent titrer miR-1a/miR-206 (les ARNm des gènes *Tmsb4x* et *Tpm4*) et cinq peuvent titrer miR-133 (les ARNm des gènes *Snap23*, *Hsd17b11*, *Eid1*, *Fam132a* et *Nptxr*). Nous avons vérifié expérimentalement l'effet du candidat le plus abondamment exprimé, *Tmsb4x* : en mutant son site d'interaction à miR-1a/miR-206 par édition du génome, nous avons effectivement observé que le miARN devenait plus actif dans la répression d'un rapporteur artificiel.

Cependant, si les ARNm extrêmement abondants peuvent parfois titrer les miARN qui les reconnaissent, il semble peu probable que ce soit le cas de toutes les cibles prédites de miARN (plus de 60% des gènes codants humains ; [Friedman *et al.*, 2009]). Il reste donc un grand nombre de gènes qui portent des sites de complémentarité conservés pour les miARN, mais qui semblent à la fois trop peu sensibles à la petite répression guidée par les miARN, et trop peu abondamment exprimés pour pouvoir efficacement moduler leur disponibilité. Nous avons donc cherché de nouvelles sources de faux positifs dans les prédictions informatiques de cibles de miARN.

Puisqu'il existe environ 2000 miARN connus chez l'Homme et chez la Souris², il est possible que des séquences génomiques soient conservées pour des raisons indépendantes des miARN, tout en présentant une complémentarité fortuite pour l'une des séquences-graines répertoriées. La graine est un hexamère, il existe donc $4^6 = 4096$ séquences-graines possibles, ce qui implique que, aussi bien chez l'Homme que chez la Souris, environ une graine possible sur deux ou trois est effectivement la graine d'au moins un miARN connu chez cette espèce. Des sites de reconnaissance de *RNA-binding proteins* dans les 3'UTR, et des sites de fixation de protéines de liaison à l'ADN (qui se trouveraient sur des 3'UTR de gènes) pourraient donc être phylogénétiquement conservées pour ces raisons indépendantes des miARN – tout en étant fortuitement complémentaires à la graine de miARN : cette coïncidence ne serait donc pas spécialement rare, et ces séquences seraient perçues par les programmes de prédiction de cibles comme des sites de fixation conservés pour le miARN, alors que leur conservation n'est pas due à l'interaction avec le miARN.

Pour estimer la fréquence de ce genre de faux positifs, nous avons mesuré la proportion de sites complémentaires à la graine de miARN connus, qui se trouvent dans des 3' UTR, et qui sont plus conservés que le miARN lui-même. C'est probablement une sous-estimation (puisque ce type de faux positifs pourrait aussi exister dans des génomes où le miARN existe également), mais nous avons malgré tout mesuré des taux de faux positifs de plusieurs dizaines de pourcent (environ 30 à 70%, selon les programmes de prédiction de cibles). Nous proposons donc que même les approches basées sur la génomique comparative capturent de nombreux faux positifs.

Suite à la publication de l'ensemble de ces résultats [Pinzón *et al.*, 2017], nous avons été sollicités par les éditeurs de deux journaux pour faire le point, dans deux articles de revue, sur les méthodes actuelles d'identification de cibles de miARN, et leurs limites [Seitz, 2017, Mockly et Seitz, 2019].

2.2 Cibles fonctionnelles du miARN *bantam* chez la Drosophile

Nos travaux montrent donc que l'identification des cibles de miARN est actuellement trop peu stringente, elle capture essentiellement des interactions moléculaires qui peuvent être bien réelles, mais sans importance biologique : les conséquences de ces événements de régulation sont atténués par l'homéostasie, et elles ne provoquent pas de changement physiologique à l'échelle macroscopique.

Parmi les myriades de cibles supposées, seule une petite fraction semble donc *fonctionnellement* ciblée, et ces rares gènes susceptibles à l'action des miARN seraient responsables des phénotypes

2. Certains miARN partagent la même séquence-graine, mais pas au point de changer les ordres de grandeur : sur la dernière version du répertoire miRBase, qui recense les miARN connus, les 2656 miARN humains partagent 1544 séquences-graines, et les 1978 miARN murins partagent 1263 séquences-graines.

macroscopiques contrôlés par les miARN. Cette conception est en opposition avec la notion couramment admise par la communauté depuis une quinzaine d'années [Bartel et Chen, 2004], selon laquelle l'action biologique des miARN passe par la régulation coordonnée de nombreux gènes, chacun étant très peu réprimé. Il faut cependant constater que, pour les interactions miARN/cibles les mieux caractérisées *in vivo*, le phénotype contrôlé par le miARN semble souvent l'être par la régulation d'une cible principale [Seitz, 2009]. Plus récemment, par une approche élégante d'édition du génome chez le Nématode *C. elegans*, il a été établi que le phénotype de létalité du mutant du miARN *let-7* semblait dû à la dérégulation d'une unique cible, *lin-41*, sans que les autres gènes réprimés par ce miARN ne semblent y contribuer [Ecsedi *et al.*, 2015].

Par une approche similaire, nous cherchons à déterminer la contribution de cibles individuelles au phénotype contrôlé par un miARN de *Drosophile*, appelé *bantam*. Ce phénotype est particulièrement spectaculaire, et facile à mesurer : *bantam* contrôle la croissance et la viabilité [Hipfner *et al.*, 2002, Brennecke *et al.*, 2003]. Après 17 ans de recherche, de nombreuses cibles ont été proposées pour ce miARN, mais sur la foi d'expériences de biologie moléculaire, et il n'a encore jamais été apporté de preuve de leur implication dans les phénotypes physiologiques. Pire : des cibles peuvent avoir été rigoureusement validées en biologie moléculaire, leur répression par le miARN est claire, elle est confirmée bio-informatiquement par sa conservation phylogénétique, mais quand on abolit leur répression *in vivo*, il s'avère qu'elles ne contribuent pas aux phénotypes physiologiques observés (comparer par exemple [Becam *et al.*, 2011] et [Bassett *et al.*, 2014]).

En mutant, par édition du génome, le miARN et les cibles de manière compensatoire (c'est à dire : le miARN muté reconnaît les cibles mutées), nous voulons abolir et restaurer sélectivement certaines interactions entre *bantam* et des cibles d'intérêt (voir Figure 1A), donc mesurer précisément la contribution d'une cible (ou : d'un ensemble de cibles) au phénotype global contrôlé par *bantam*.

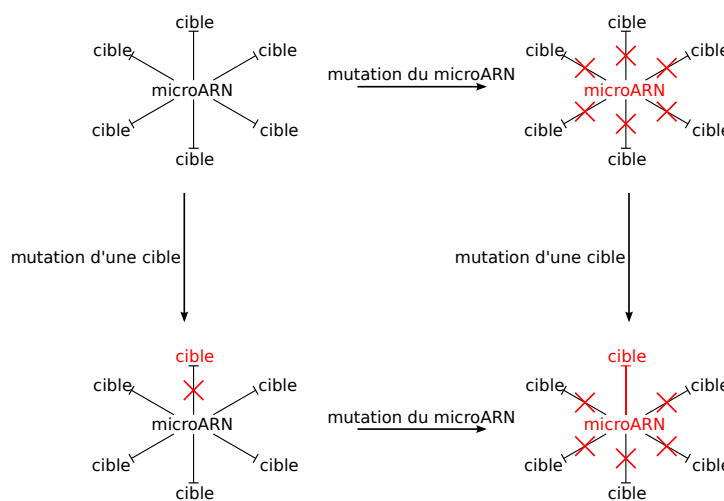


FIGURE 1 – **Quantification des conséquences physiologiques d'interactions individuelles entre *bantam* et des cibles d'intérêt.** **A.** Muter le miARN seul (en haut à droite) abolit les interactions entre le miARN et toutes ses cibles, révélant le phénotype global contrôlé par le miARN. Muter son site de fixation sur une cible seule (en bas à gauche) abolit uniquement son interaction avec la cible désirée, révélant la contribution de cette interaction au phénotype global. Muter simultanément le miARN et son site de fixation sur une cible, de manière compensatoire (en bas à droite), restaure uniquement cette interaction et abolit toutes les autres, révélant la contribution de toutes les autres cibles (à l'exclusion de la cible d'intérêt) au phénotype global. **B.** Les éventuels phénotypes observés quand un site d'interaction au miARN sont mutés peuvent être dus à la perte d'interaction avec le miARN, ou à la perte d'interaction avec un autre facteur (dont le site de fixation serait fortuitement complémentaire au miARN). La mutation compensatoire du miARN restaurerait le phénotype sauvage dans le 1er cas (scénario (i)), mais pas dans le 2ème (scénario (ii)).

Une expérience de mutation compensatoire entre un miARN et une cible a donc déjà été réalisée chez le Nématode, et elle a montré que le phénotype de létalité dû au miARN *let-7* était essentiellement attribuable à une unique cible, *lin-41* (muter de façon compensatoire le miARN et la cible restaure la viabilité du ver, alors que ce miARN muté ne reconnaît aucune des nombreuses autres cibles prédites; [Ecsedi *et al.*, 2015]). Mais l'interprétation quantitative de ces résultats est compliquée par la nature de la mutation du miARN *let-7* utilisée : ce miARN muté n'est pas mûri correctement, il est sous-accumulé, or la viabilité du ver n'est que partiellement restaurée. Il est donc impossible de savoir si le sauvetage n'est que partiel parce que d'autres cibles contribuent marginalement à la létalité, ou parce que le miARN muté est trop peu abondant pour réprimer correctement la cible mutée [Ecsedi *et al.*, 2015].

Dans nos propres expériences, nous apportons donc un soin particulier au contrôle de l'expression du miARN muté. Nous avons préparé trois souches différentes, où la graine du miARN est mutée ($\text{GAGAUC} \rightarrow \text{CAAAAU}$; l'hexanucléotide CAAAAU ne constitue la graine d'aucun miARN connu chez la Drosophile) tout en préservant la structure secondaire du précurseur en tige-boucle du miARN (cette structure est nécessaire à la maturation du miARN). Les trois souches diffèrent par des mutations supplémentaires, qui pourraient affecter la biogenèse du miARN (l'allèle « m22 » ne contient pas d'autre mutation dans la tige-boucle du pré-miARN, mais il porte une mutation à 11 pb du pied de la tige-boucle, dans le PAM que nous avons utilisé pour le clivage par Cas9; l'allèle « m27 » ne contient pas d'autre mutation dans la tige-boucle du pré-miARN, mais il porte une mutation sur un PAM plus éloigné, à 54 pb du pied de la tige-boucle; et l'allèle « m47 » porte la même mutation sur le PAM éloigné, ainsi qu'une mutation dans le brin *bantam*^{*} de la tige-boucle, qui devrait augmenter l'asymétrie du duplex *bantam*/*bantam*^{*} et donc l'abondance du miARN; [Schwarz *et al.*, 2003, Khvorova *et al.*, 2003]). Outre ces trois allèles, nous avons préparé un quatrième allèle mutant (allèle « md »), qui délète une grande partie de la tige-boucle pré-*bantam*, et qui ne génère aucun petit ARN : cet allèle servira de contrôle, dont les phénotypes ne devraient pas être sauvés par d'éventuelles mutations compensatoires sur les cibles (voir Figure 2).

Ces souches sont maintenant établies, et nous avons constaté que chacun de nos mutants récapitule, à l'état homozygote, le phénotype de défaut de croissance et de létalité larvaire et pupale décrits pour le mutant nul historique [Hipfner *et al.*, 2002]. Nous nous apprêtons maintenant à mesurer l'expression du miARN muté, *in vivo* au 1^{er} stade larvaire (stade développemental où l'expression de *bantam* est élevée : [Brennecke *et al.*, 2003, Leaman *et al.*, 2005]); pour des raisons techniques, il est beaucoup plus facile de purifier de grandes quantités de larves hétérozygotes que d'embryons hétérozygotes). Parmi les trois souches exprimant un *bantam* muté (les souches appelées « m22 », « m27 » et « m47 »), nous choisirons celle qui produit principalement le miARN souhaité, avec les bonnes extrémités 5' et 3', et dont le niveau d'expression est, quantitativement, la plus proche de celui de *bantam* sauvage. Ces mesures seront effectuées par Small RNA-Seq sur des hétérozygotes, de manière à utiliser leur allèle sauvage comme contrôle interne. Nous avons également vérifié que le miARN muté ne provoquait pas des phénotypes par le simple fait de son expression (il aurait pu fortuitement se trouver complémentaires de gènes dont la répression provoquerait des phénotypes visibles) : nous n'avons observé aucun défaut particulier chez les mouches transgéniques qui expriment cette tige-boucle mutée sous le contrôle du promoteur ubiquitaire du gène *daughterless*.

Par ailleurs, nous avons sélectionné plusieurs cibles prédites de *bantam*, dont nous avons commencé à muter les sites de fixation de *bantam* de manière compensatoire à notre mutation du miARN (les 3' UTR de ces gènes contiennent l'hexamère GAUCUC, complémentaire à la graine de *bantam* sauvage; nous les mutons en AUUUUG, complémentaire à la graine de *bantam* muté). Voici les 7 cibles prédites que nous avons décidé de muter sur leurs sites de fixation du miARN :

- Parmi les gènes qui portent des sites de reconnaissance de la graine de *bantam* phylogénétiquement conservés, le gène *hid* et le gène *tutl* ont un statut particulier : leurs 3' UTR contiennent respectivement 5 et 3 sites de reconnaissance conservés pour *bantam*. Toutes les autres cibles prédites de *bantam* ont, au maximum, 1 site conservé. Cette particularité suggère que *hid* et *tutl* sont plus fortement réprimés [Doench *et al.*, 2003], et elle a rapidement suggéré que

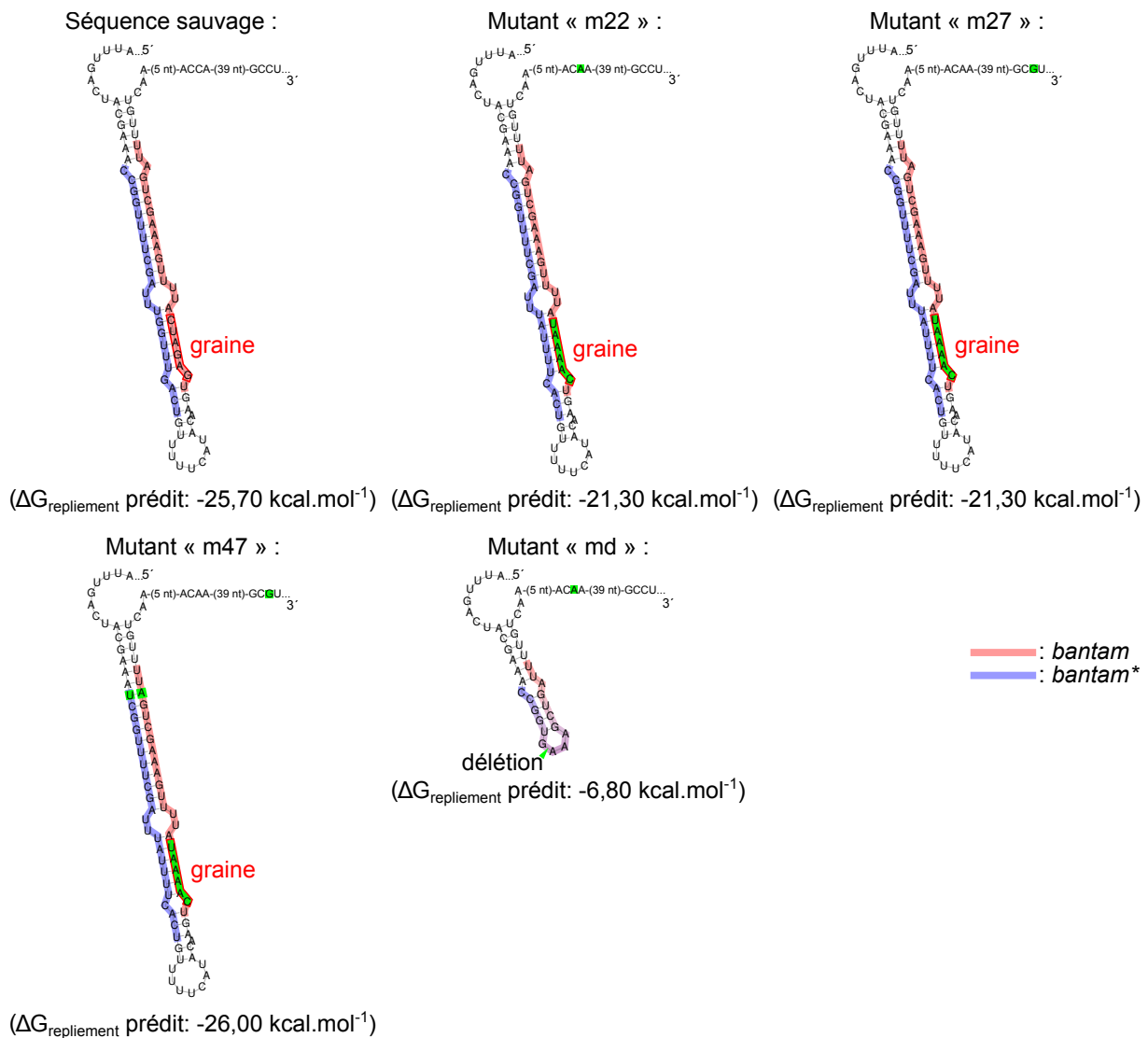


FIGURE 2 – **Série allélique *bantam***. La structure secondaire des tiges-boucles, et son enthalpie libre de repliement, ont été prédites par le programme **RNAfold**. Les mutations qui distinguent les mutants de la séquence sauvage sont indiquées en vert. La séquence mature du miARN *bantam* est surlignée en rouge, celle du miARN* *bantam** en bleu, et la graine du miARN est encadrée en rouge.

hid était une cible fonctionnellement importante pour *bantam* [Brennecke *et al.*, 2003]. Mais une validation rigoureuse de cette importance phénotypique n'a jamais été apportée. Nous avons donc inclus *hid* et *tutl* dans les cibles à étudier : que nos expériences confirment leur importance fonctionnelle ou non, il s'agira d'un grand progrès.

- Il est connu que l'effet de *bantam* sur la croissance était dû, au moins partiellement, à des défauts dans la voie de l'hormone ecdysone [Boulan *et al.*, 2013]. Le contrôle de la croissance et de la chronologie du développement par *bantam* est conservé entre *Drosophila melanogaster* et *Cænorhabditis elegans* [Boulan *et al.*, 2013] : nous nous intéresserons donc aux gènes dont l'interaction avec *bantam* est conservée entre ces deux espèces. Nos analyses informatiques préliminaires montrent que 5 gènes impliqués dans la production ou la sensibilité à l'ecdysone ont des sites de fixation pour *bantam* conservés entre le Nématode et la Drosophile : nous incluons donc ces 5 gènes dans notre étude.

Pour l'heure, nous avons établi l'une de ces 7 souches : la souche mutée sur le gène *hid* (chacun de ses 5 sites de reconnaissance de *bantam* est muté; ce mutant est appelé *hid^{mut5}*). Ces mouches sont viables et fertiles à l'état homozygote, et nous n'avons observé qu'un seul phénotype remarquable :

un retard de la métamorphose (voir Figure 3).

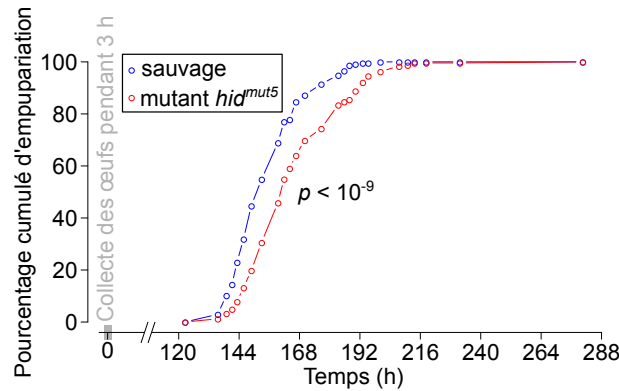


FIGURE 3 – **Retard développemental du mutant *hid^{mut5}***. Pour chaque génotype, nous avons préparé 10 tubes contenant, chacun, une trentaine de larves L1 (collectées pendant une période de 3 h). La distribution cumulée du nombre d'empupariations (normalisée par le nombre total de pupes accumulées en fin d'expérience) montre que les larves mutées sur les 5 sites de fixation de *bantam* sur *hid* s'empupent environ 10 h plus tard que les larves sauvages. La significativité de la différence a été mesurée par un test de permutation à partir de la distribution empirique.

L'abolition de l'interaction entre le miARN *bantam* et le gène *hid* provoque donc un phénotype beaucoup moins sévère que la perte du miARN *bantam* : les phénotypes de défaut de croissance et de létalité du mutant *bantam* ne sont donc pas uniquement dus à la dérégulation de *hid*, d'autres cibles doivent y contribuer. Nous entamons l'analyse de mutants sur d'autres cibles prédites.

2.3 Contrôle de la prolifération cellulaire par miR-34

Dans le but également de mesurer la contribution de cibles individuelles au phénotype contrôlé par un miARN, nous avons imaginé une approche de crible à grande échelle sur des cellules en culture : si un miARN réprime la prolifération cellulaire, la mutation des sites de fixation du miARN responsables de ce phénotype devrait accélérer la prolifération. Donc dans une population cellulaire, si certaines cellules sont mutées sur un tel site fonctionnel, les clones issus de ces cellules devraient se sur-accumuler, et une mesure de cette accumulation par séquençage à haut débit devrait identifier les cibles dont la répression par le miARN ralentirait la prolifération.

Nous avons donc choisi d'étudier les miARN miR-34 dans des cellules de Mammifères : cette famille de miARN³ est activée transcriptionnellement par le facteur de transcription p53, qui est surexprimé à la suite de stress oncogéniques [Raver-Shapira *et al.*, 2007, Chang *et al.*, 2007, He *et al.*, 2007]. La réponse p53 a un effet anti-prolifératif, et la sur-expression expérimentale des miARN miR-34 dans diverses lignées cellulaires humaines et murines réprime également la prolifération, ce qui a laissé penser que miR-34 participait à l'action de p53 sur la prolifération [Raver-Shapira *et al.*, 2007, Chang *et al.*, 2007, He *et al.*, 2007].

Puisque cet effet anti-prolifératif de miR-34 a été observé à la fois dans des cellules humaines et murines, nous avons fait l'hypothèse que les interactions miR-34/cibles responsables de l'effet anti-prolifératif étaient, elles-mêmes, conservées entre ces deux espèces. Nos analyses ont identifié 1420 sites de complémentarité à la graine de miR-34, situés dans des 3' UTR, et conservés entre Homme et Souris. Nous avons donc envisagé de préparer une banque CRISPR pour cibler ces 1420 sites (ainsi que 500 sites-contrôles, non reconnus par miR-34), par 3 sgRNA par site. Notre objectif était ensuite

3. On appelle « famille de miARN » l'ensemble des miARN qui partagent la même séquence-cible. Puisque la graine est le principal déterminant de la reconnaissance des cibles, les membres d'une même famille ciblent essentiellement les mêmes gènes [Bartel, 2009].

de co-transduire cette banque, avec Cas9, dans une culture de cellules humaines HCT-116 (une lignée essentiellement diploïde, ce qui devait faciliter l'obtention de clones homozygotes mutants). En adaptant la multiplicité d'infection de manière à ce qu'une cellule soit transduite, le plus souvent, par 1 et 1 seul sgRNA, nous comptons donc préparer une population hétérogène, constituée de clones, chacun muté aléatoirement sur un site de fixation de miR-34 dans un gène donné. Les clones mutés sur un site qui contribue à la répression de prolifération par miR-34 devraient donc proliférer plus vite, et s'accumuler dans la population cellulaire ; un séquençage à haut débit de l'ADN génomique de la population (pas un séquençage du génome entier, mais d'un enrichissement pour les 1420+500 sites ciblés, de manière à garantir une plus grande couverture de séquençage pour la même profondeur totale) allait donc nous permettre d'identifier ces sites, et en même temps, de mesurer leur effet anti-prolifératif individuel (voir Figure 4).

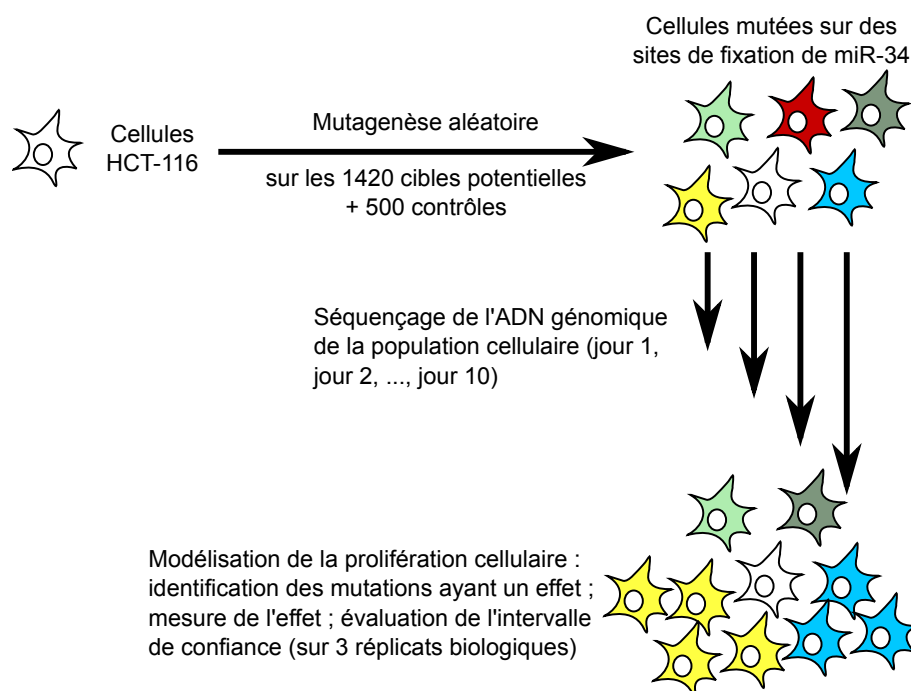


FIGURE 4 – Crible pour l'identification des cibles de miR-34 responsables de son effet anti-prolifératif. La transduction de la banque CRISPR doit aboutir à une population hétérogène, où chaque clone est muté sur un site individuel de fixation de miR-34. Après 10 jours de culture, les clones mutés sur les sites à effet anti-prolifératif devraient donc s'accumuler dans la population (par exemple, les clones représentés en jaune et en bleu sur cette figure). Le séquençage ciblé des loci visés par la banque CRISPR révélera ces changements d'abondance, et mesurera leur ampleur.

Mais nos expériences préliminaires ne confirment pas l'effet anti-prolifératif de miR-34 : lorsque nous inhibons miR-34 par transfection d'un oligonucléotide complémentaire, les cellules HCT-116 ne prolifèrent pas davantage (elles prolifèrent même un peu moins vite que les cellules traitées par un oligo contrôle ; voir Figure 5A). L'interprétation de cette expérience est compliquée par le fait que l'oligo transfecté au jour 0 est progressivement perdu par dilution au fil des divisions cellulaires. Nous avons donc également préparé des cellules HCT-116 homozygotes mutantes pour le gène *miR-34a* par édition du génome⁴, et obtenu 2 clones indépendants. Chacun prolifère plus lentement que les cellules HCT-116 parentales (voir Figure 5B). Enfin, quand nous sur-exprimons miR-34 par transfection transitoire à 100 nM, nous reproduisons l'effet anti-prolifératif qui avait été publié par [Chang *et al.*, 2007] et [He *et al.*, 2007] sur cette lignée cellulaire (voir Figure 5C), mais cet effet

4. *miR-34a* est le gène de la famille miR-34 le plus abondamment exprimé : il constitue entre 78,7 et 99,5% des reads de la famille miR-34 dans les jeux de données SRA n°SRR4235725, SRR4235726, SRR954987 et SRR954996.

disparaît si le miARN est transfecté moins abondamment, à 10 nM, une concentration plus proche de la concentration intracellulaire naturelle (voir Figure 5D).

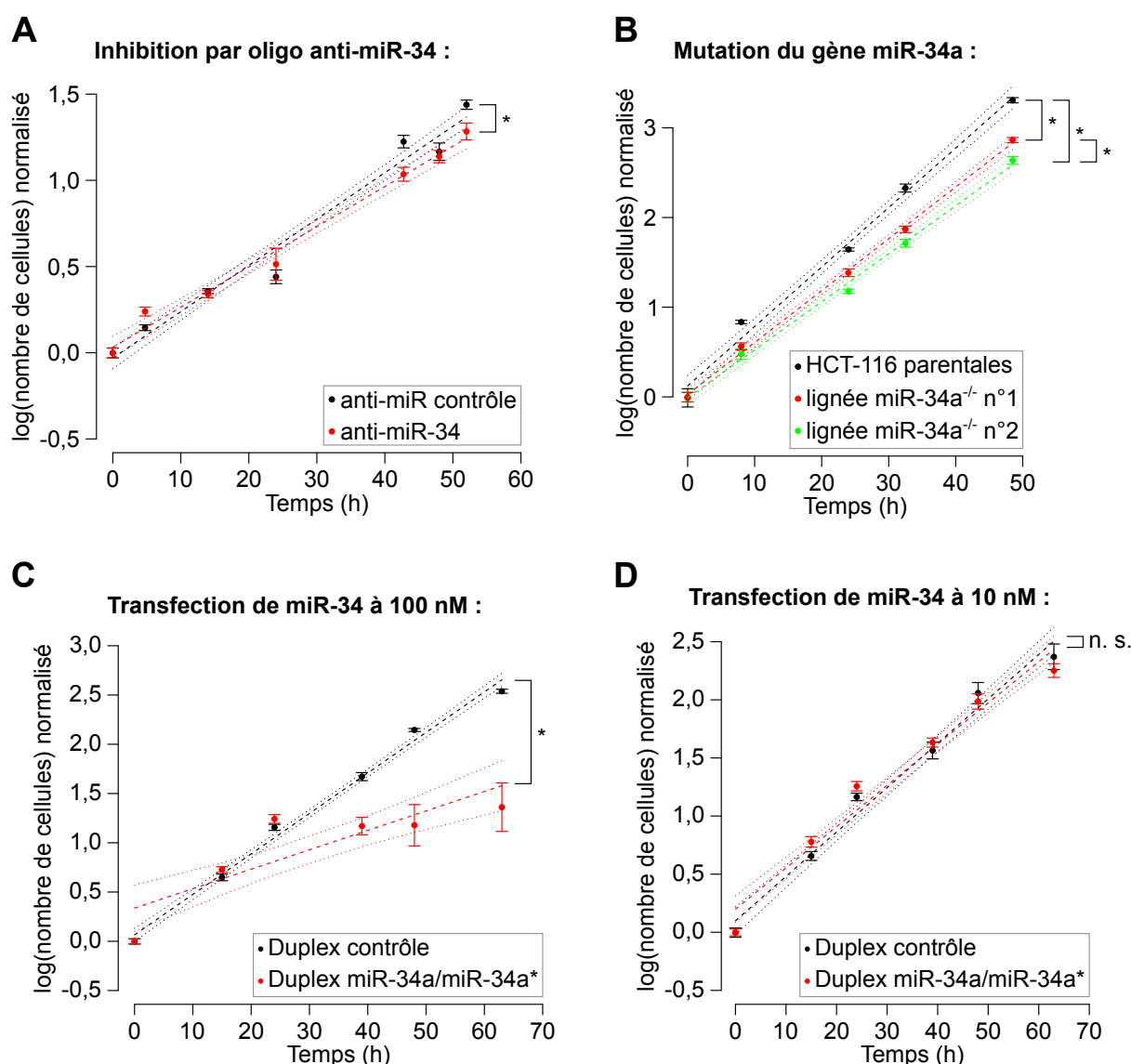


FIGURE 5 – Vérification de l'effet de miR-34 sur la prolifération des cellules HCT-116. **A.** La transfection d'un oligo 2'-O-méthylé anti-miR-34 (en rouge) à 133 nM ralentit légèrement la croissance des cellules HCT-116 par rapport à un oligo contrôle (en noir). **B.** Deux clones *miR-34a*^{-/-} indépendants (en rouge et en vert) prolifèrent légèrement moins vite que les cellules HCT-116 parentales (en noir). **C.** La transfection de 100 nM de duplex miR-34a/miR-34a* synthétique (en rouge) ralentit fortement la croissance de cellules HCT-116, par rapport à la transfection d'un duplex contrôle (en noir). **D.** La transfection de 10 nM de duplex miR-34a/miR-34a* synthétique (en rouge) n'affecte pas notablement la croissance de cellules HCT-116, par rapport à la transfection du duplex contrôle (en noir). Dans chaque panel, le nombre de cellules a été estimé par un kit de mesure de l'ATP (en quantité supposée proportionnelle au nombre de cellules), basée sur une réaction de bioluminescence : le comptage cellulaire est donc exprimé en unités arbitraires. Lignes en tirets : régression linéaire ; lignes en pointillés : limites de l'intervalle de confiance à 95% sur la régression ; barres d'erreur : erreur-type. Chaque expérience a été réalisée en 6 à 8 réplicats biologiques. « * » : pentes des régressions significativement différentes ($p - value < 0,05$) ; « n. s. » : pentes non significativement différentes.

Il semble donc que l'effet anti-prolifératif longtemps attribué à miR-34 soit finalement artefactuel,

au moins dans les cellules HCT-116 (qui faisaient partie des cellules étudiées par [Chang *et al.*, 2007] et par [He *et al.*, 2007]) : ces conclusions étaient basées sur des expériences de sur-expression du miARN (une transfection d'ARN synthétiques imitant le précurseur pre-miR-34a [Chang *et al.*, 2007] ou le duplex naturel miARN/miARN* [He *et al.*, 2007]), à des doses supra-physiologiques (60 nM pour [Chang *et al.*, 2007] et 100 nM pour [He *et al.*, 2007], alors que les concentrations intracellulaires de miARN naturels atteignent rarement une telle valeur – elles restent habituellement de l'ordre du nanomolaire; [Bosson *et al.*, 2014, Pinzón *et al.*, 2017]). Or la surexpression artificielle de miARN provoque des dérégulations artefactuelles, de diverses origines, qui ne sont pas dues à la répression des cibles naturellement régulées par des concentrations physiologiques du miARN [Mockly et Seitz, 2019].

Il se trouve que miR-34 a suscité énormément d'intérêt de la part de la communauté des oncologues moléculaires, parce que c'était le premier miARN suspecté d'avoir une action anti-cancéreuse. Des centaines d'articles lui ont été consacrés, pour tâcher d'élucider son mode d'action (mais, bien souvent, en reproduisant des sur-expressions déraisonnables), et les applications thérapeutiques de ces connaissances ont toutes échoué jusqu'à présent (c'est l'objet d'un de nos projets pour les années à venir). Il nous semble donc important d'alerter la communauté sur cet artefact probable.

2.4 Évolution de la machinerie de l'interférence à ARN chez les Animaux

Le contenu de cette section a été publié [Pinzón et al., 2019] : le lecteur est renvoyé à cet article pour les détails, et nous ne donnerons ici qu'un résumé synthétique.

L'interférence à ARN (« RNAi ») est un mécanisme de répression post-transcriptionnel initié par un ARN double-brin, et qui réprime les ARN présentant une homologie de séquence avec l'un ou l'autre des deux brins. Observé dans la plupart des clades eucaryotiques étudiés (Plantes, Champignons, Animaux, Ciliés, ...), il implique des acteurs très conservés : la ribonucléase Dicer (qui clive l'ARN double-brin déclencheur en duplex d'ARN d'une vingtaine de paires de bases, appelés « siARN »); et les protéines effectrices de la sous-famille Ago, de la famille Argonaute, où sont chargés les siARN [Ghildiyal et Zamore, 2009]. Parmi les participants à ce mécanisme, une famille de protéine a un statut particulier : les polymérase à ARN dépendant de l'ARN (RdRP, pour : *RNA-dependent RNA polymerases*). Chez de nombreuses espèces, elles sont requises pour le RNAi : le RNAi est aboli chez des mutants de ces enzymes chez des Champignons [Cogoni et Macino, 1999], des Plantes [Mourrain *et al.*, 2000, Dalmay *et al.*, 2000], et des Ciliés [Marker *et al.*, 2010]. Les Animaux semblent faire exception, puisqu'un seul organisme-modèle animal est connu pour utiliser une RdRP pour l'interférence à ARN : le Nématode [Smardon *et al.*, 2000, Sijen *et al.*, 2001]. Les autres organismes-modèles animaux courants (Vertébrés et Drosophile) ne possèdent d'ailleurs même pas de gène de RdRP, alors qu'ils sont en mesure de réaliser le RNAi [Ghildiyal et Zamore, 2009].

Non seulement l'implication des RdRP est donc variable dans les différents clades eucaryotiques, mais leur mécanisme d'action également : chez les Plantes et les Champignons, on considère que les RdRP polymérisent de longs ARN antisens à partir d'un ARN-matrice, et ces deux ARN, complémentaires, doivent s'hybrider, et former un ARN double-brin qui servira de substrat à Dicer [Allshire, 2002, Motamedi *et al.*, 2004, Voinnet, 2008, Martienssen et Moazed, 2015]. Chez le Nématode en revanche, des RdRP très peu processives ne fabriquent que de petits ARN longs d'une vingtaine de nucléotides : ces petits ARN simple-brin sont ensuite directement chargés sur une protéine Argonaute, sans implication de Dicer, et sans intermédiaire double-brin [Pak et Fire, 2007, Sijen *et al.*, 2007, Gu *et al.*, 2009]⁵. Alors que les siARN, fabriqués par Dicer, portent un monophosphate en 5' (puisque Dicer clive son substrat ARN double-brin sur un pont phosphodiester, en laissant le phosphate sur son produit 3'), les produits bruts de la polymérisation par une RdRP portent un 5' triphosphate (le premier nucléotide polymérisé conserve ses trois phosphates, alors

5. Notez que l'intégrité scientifique de la référence [Sijen *et al.*, 2007] a été sérieusement mise en question : <https://pubpeer.com/publications/2B00E5BEB5B75B499550D03C15EFA4>.

que les suivants ne conservent que leur phosphate α au cours de la polymérisation). Parmi les petits ARN produits par les RdRP de Nématode⁶, on distingue deux classes : les « 22G » (qui mesurent 22 nt de long, et portent un G en position 5' pour la presque totalité d'entre eux), et les « 26G » (qui mesurent 26 nt de long, et portent un G en position 5'). Les 22G portent un 5' triphosphate [Pak et Fire, 2007], probablement parce que leur extrémité 5' ne subit aucune maturation après polymérisation par la RdRP. Les 26G, en revanche, portent un 5' monophosphate, et on ne sait pas s'ils dérivent de la maturation d'ARN initialement triphosphorylés, ou s'ils sont directement polymérisés sous la forme d'ARN monophosphorylés (par exemple, en amorçant leur polymérisation par un nucléotide monophosphate) [Ruby *et al.*, 2006, Han *et al.*, 2009, Vasale *et al.*, 2010].

Nous avons décidé de tirer parti de la disponibilité de nombreux génomes séquencés chez les Métazoaires (et donc, de nombreux protéomes prédits) pour rechercher des RdRP parmi les protéomes prédits d'un grand nombre d'espèces. Il existe 20 familles de RdRP annotées dans la banque de données PFAM (version 31 au moment de notre analyse) : 19 de ces 20 familles sont des familles de protéines virales, qui réalisent la réplication ou la transcription de génomes ARN viraux, et ne présentent pas de similarité de séquence avec les RdRP impliquées dans le RNAi. Toutes les RdRP connues pour participer au RNAi appartiennent à la 20ème famille, appelée « RdRP eucaryotiques » (PFAM n°PF05183 ; c'est la seule famille de RdRP non virales), et l'amplification du RNAi constitue la seule fonction connue de cette famille de protéines.

En compilant les différentes banques de données génomiques, nous avons pu rassembler les protéomes prédits de 538 espèces animales. Une recherche de protéines de la famille des RdRP eucaryotiques, par **HMMer**⁷, identifie des RdRP eucaryotiques dans une grande variété d'espèces (Mollusques, Chélicérates, ...), notamment dans un clade où on les avait toujours pensées absentes (les Insectes ; [Lewis *et al.*, 2018, Li *et al.*, 2018]) : voir Figure 6.

Une distribution aussi sporadique pourrait être due à de multiples transferts horizontaux, ou à une origine commune, suivie de nombreuses pertes indépendantes, dans des lignages récents. La comparaison des séquences de toutes ces RdRP suggère que les transferts horizontaux ont été très rares : la phylogénie des RdRP tend à suivre celle des Métazoaires eux-mêmes, ce qui indique très vraisemblablement une origine commune, et de nombreuses pertes récentes.

Afin d'évaluer la conservation de la fonction de ces RdRP dans l'amplification du RNAi, nous avons choisi une espèce animale possédant des RdRP apparemment intactes, qui serait phylogénétiquement éloignée du Nématode, et dont nous pourrions obtenir simplement des échantillons de divers stades développementaux, notamment des adultes prélevés dans leur milieu naturel (puisque le RNAi joue souvent un rôle de défense contre les pathogènes présents dans le milieu naturel [Guo *et al.*, 2019], et qu'il est concevable qu'il puisse également réguler les transitions développementales). Notre choix s'est porté sur un Céphalochordé, *Branchiostoma lanceolatum* (indiqué en Figure 6), un organisme-modèle émergent dont le génome et l'épigénome étaient en cours de séquençage et annotation par le consortium AmphienCODE, sous la coordination d'Hector Escriva (UMR 7232, à Banyuls).

Nous avons séquencé les petits ARN (de 18 à 30 nt) à partir de plusieurs stades embryonnaires, juvénile et adulte, en prenant soin d'utiliser des protocoles expérimentaux qui nous permettraient de détecter aussi bien les petits ARN 5' mono- et poly-phosphorylés, et de distinguer les ARN portant un 2'-OH et 3'-OH des ARN chimiquement modifiés sur leur 3' (par exemple : porteurs d'une 2'-O-méthylation en 3', comme c'est le cas pour les siARN de Drosophile : [Pélisson *et al.*, 2007]). Alors que les autres classes de petits ARN (miARN et piARN) étaient facilement observés, nous

6. Ces petits ARN sont appelés « siARN secondaires », parce que leur biogenèse est déclenchée par les siARN produits par Dicer : c'est la reconnaissance d'un ARN-cible par un siARN qui recrute la RdRP sur cet ARN, et induit donc la polymérisation de siARN secondaires complémentaires de cet ARN [Pak et Fire, 2007, Sijen *et al.*, 2007, Gu *et al.*, 2009].

7. Le programme **HMMer** tient compte des importances relatives de chaque position dans l'alignement, en distinguant les acides aminés peu conservés des acides aminés très conservés (comme ceux trouvés dans les sites catalytiques par exemple), à la différence de **blast** (qui ne pondère pas les différentes positions) : **HMMer** est donc beaucoup plus sensible, spécialement pour des séquences très divergentes.

n'avons pas pu trouver de siARN chez *B. lanceolatum*. Même en multipliant les analyses, pour trouver d'éventuels petits ARN qui auraient pu passer inaperçus (potentiels siARN dirigés contre des génomes de pathogènes, et qui ne présenteraient d'homologie ni au génome ni au transcriptome de *B. lanceolatum*; ARN antisens peu abondants, dilués dans les populations de miARN et de piARN; ARN présentant une homologie à des génomes de pathogènes connus de *Branchiostoma*), nous n'avons jamais observé de population de petits ARN présentant les caractéristiques attendues des siARN (taille stéréotypée, biais de séquence sur la position 5').

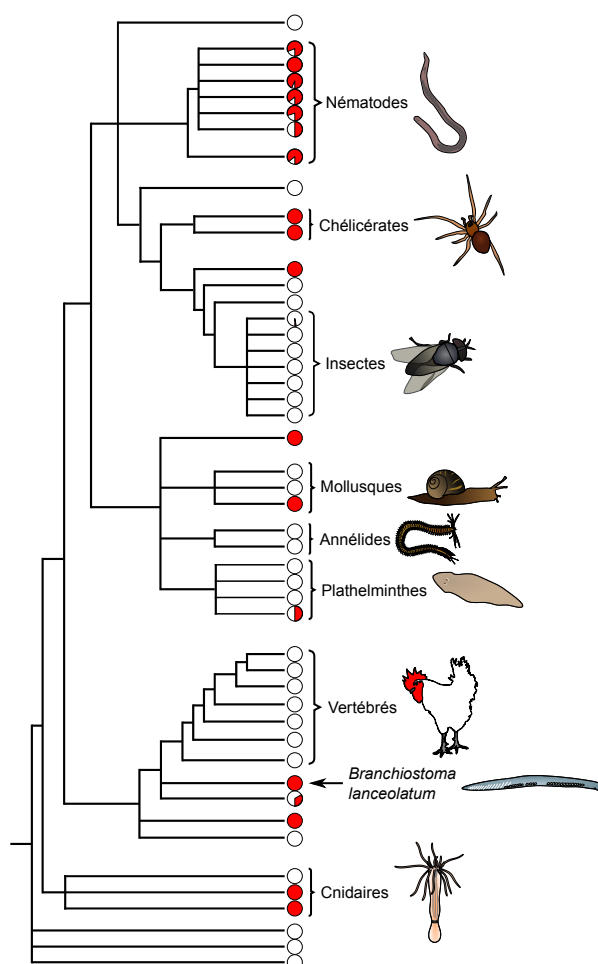


FIGURE 6 – **Distribution phylogénétique sporadique des RdRP eucaryotiques chez les Animaux.** Pour chaque taxon analysé, le secteur rouge du diagramme circulaire représente la proportion d'espèces de ce taxon dont le génome contient au moins un gène de RdRP eucaryotique. La distribution de cette famille de gènes est très sporadique chez les Animaux, ce qui indique qu'elle a été perdue récemment dans de nombreux lignages.

Nous concluons donc que les RdRP de *B. lanceolatum* ne participent pas à l'amplification du RNAi. Cette fonction était la seule connue pour les RdRP eucaryotiques : la conservation de ces gènes (et notamment, pour trois loci de *B. lanceolatum* : la conservation de leur site catalytique) suggère cependant qu'elles jouent un rôle biologique dans cette espèce. Nous n'avons pas idée de ce que pourrait être cette fonction, indépendante du RNAi. Nous remarquons cependant que le Champignon *Aspergillus nidulans*, qui possède deux gènes de RdRP avec des sites actifs intacts, n'en utilise aucun pour le RNAi [Hammond et Keller, 2005].

Outre la recherche de l'activité des RdRP, et par cette analyse des petits ARN de *B. lanceolatum*, nous avons également participé à l'annotation de son génome, publié quelques mois plus tôt [Marlétaz *et al.*, 2018].

2.5 Recherche d'inducteurs de la dégradation des microARN

Alors que le contrôle de la biogenèse des miARN (transcription et maturation) fait l'objet d'études dédiées depuis une quinzaine d'années [O'Donnell *et al.*, 2005, Viswanathan *et al.*, 2008], la régulation de leur stabilité a longtemps été ignorée. Pourtant, des observations suggéraient que certains miARN étaient soumis à une dégradation active, au moins dans les neurones optiques [Krol *et al.*, 2010]. Très récemment, des mesures à haut débit, dans diverses lignées cellulaires, ont confirmé que les demi-vies des miARN étaient très variables (de quelques heures à quelques jours; [Reichholf *et al.*, 2019, Kingston et Bartel, 2019]), suggérant à nouveau une dégradation régulée.

Mais c'est par sérendipité que les premiers régulateurs de stabilité des miARN ont été découverts : dans chaque exemple connu, il s'agit d'ARN reconnus par le miARN (comme le seraient des ARN-cibles habituellement), et qui induisent la dégradation exonucléolytique du miARN, de 3' vers 5' ; bien souvent, des isoformes allongées (qui présentent des additions de nucléotides en 3', principalement des uridines, plus rarement des adénosines) sont également observées : il y a une controverse pour savoir si ces additions de nucléotides constituent un signal qui dirige le miARN vers la voie de dégradation, ou s'il s'agit de produits secondaires [Ameres *et al.*, 2010, Cazalla *et al.*, 2010, Marciniowski *et al.*, 2012, de la Mata *et al.*, 2015, Bitetti *et al.*, 2018, Kleaveland *et al.*, 2018].

Ce processus a été baptisé TDMD (*Target RNA-directed microRNA degradation*). Avec l'accumulation d'exemples d'ARN inducteurs de TDMD, *in vivo* et *ex vivo*, commencent à émerger des motifs communs : les appariements ARN long/miARN qui déclenchent le TDMD tendent à impliquer un long appariement ininterrompu sur l'extrémité 3' du miARN (une telle géométrie est très rare dans les interactions miARN/cibles habituelles) [Fuchs Wightman *et al.*, 2018].

À partir d'une compilation d'appariements connus pour induire ou non le TDMD, nous avons ainsi pu formaliser les règles de prédiction de fonctionnalité des appariements. Nous avons ensuite cherché des sites susceptibles de s'apparier avec les miARN selon ces géométries, dans les 100 génomes de Vertébrés utilisés pour l'alignement pan-génomique multiple, centré sur le génome humain, de l'*UCSC Genome Browser* [Rosenbloom *et al.*, 2015]. Cette recherche nous a ensuite permis de mesurer, pour chaque site, sa conservation phylogénétique parmi les Vertébrés (*N.B.* : nous avons exclu les sites localisés dans des séquences codantes annotées, dont la conservation pourrait s'expliquer trivialement par la pression de sélection sur la séquence protéique codée).

Pour sélectionner des candidats à tester expérimentalement, nous avons dû choisir un système cellulaire. Il semble que (pour des raisons encore incomprises) le TDMD soit particulièrement actif dans les neurones [Fuchs Wightman *et al.*, 2018]. Nous avons donc décidé de réaliser ces expériences dans les neurones corticaux de Souris : c'est un type cellulaire où le TDMD a déjà été mis en évidence [de la Mata *et al.*, 2015], et il est très abondant dans chaque souris (ce qui permet de limiter le nombre de sacrifices animaux pour obtenir la quantité de matériel biologique nécessaire).

Nous avons pu évaluer l'abondance de chaque ARN-candidat dans les neurones corticaux de Souris par des données de RNA-Seq publiques. La Figure 7 représente (en rouge) les 4 candidats que nous avons décidé de tester expérimentalement. Parmi les deux exemples d'inducteurs de TDMD caractérisés *in vivo* (les ARNm Nrep et Cyrano; [Bitetti *et al.*, 2018, Kleaveland *et al.*, 2018]), nous avons décidé d'inclure Nrep dans nos tests de validation – mais Cyrano était également détecté par notre approche (voir Figure 7). Chacun des 4 candidats sélectionnés est un ARNm, dont le site potentiellement inducteur de TDMD se trouve dans sa 3' UTR.

En collaboration avec Énora Moutin et Julie Perroy (UMR 5203, à Montpellier), qui nous ont fourni des neurones corticaux de Souris, nous avons réprimé ces 4 candidats par RNAi (deux réplicats biologiques par candidat), en transduisant ces neurones avec des vecteurs lentiviraux exprimant, chacun, 3 à 4 shRNA contre le gène d'intérêt. Nous avons extrait les ARN de ces cellules 10 jours après la transduction, et soumis les échantillons à la plateforme de séquençage MGX, pour préparation et séquençage des 8 banques de Small RNA-Seq. Nous sommes en attente des résultats.

Par ailleurs, et dans la perspective d'une analyse plus poussée des déterminants du TDMD, que l'activité TDMD du candidat *Nnat* soit validée ou non, nous avons commencé à préparer des

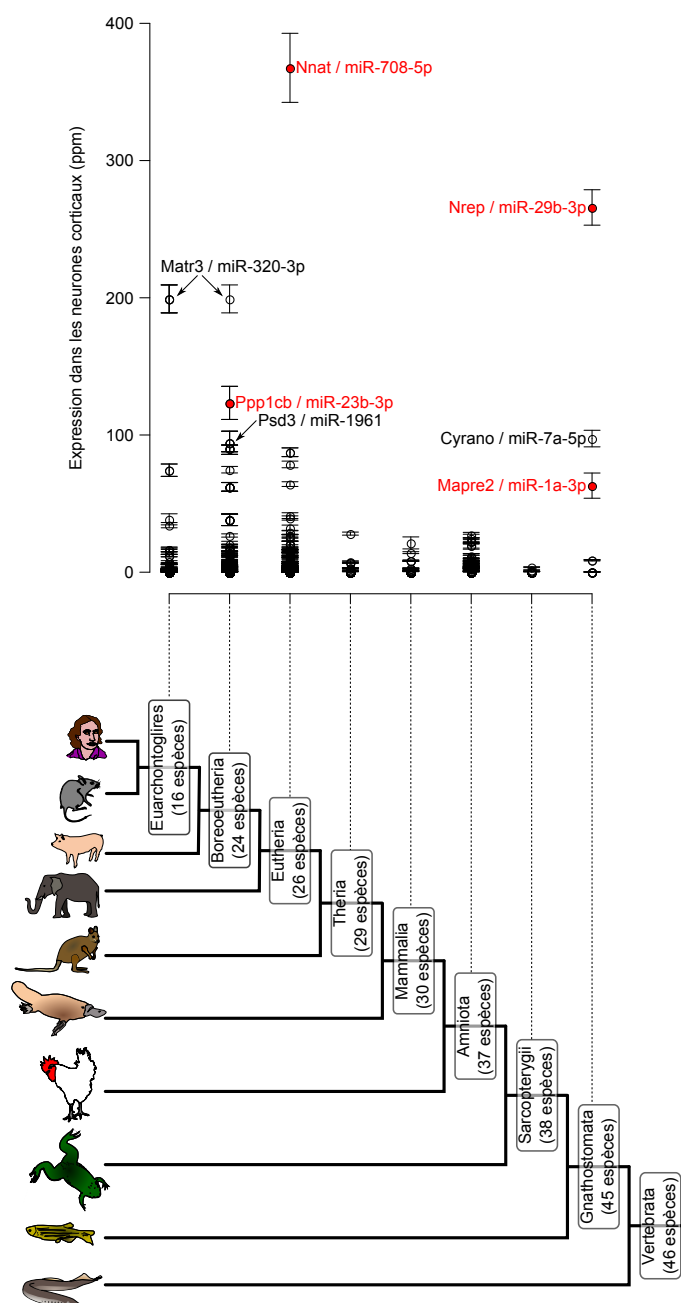


FIGURE 7 – Conservation des candidats inducteurs de TDMD, et abondance dans les neurones corticaux de Souris. Sont représentés les candidats spécifiques de l'un des clades représentés sous le graphique (un candidat est considéré spécifique d'un clade s'il est conservé dans au moins 50% des espèces de ce clade, tout en étant absent de chacune des espèces extérieures au clade ; notez que cette procédure exclut des candidats, dont la spécificité de clade ne peut être établie clairement : ce sont les candidats qui ne sont pas conservés dans au moins 50% des espèces du plus grand clade contenant toutes les espèces possédant ce candidat). L'expression de chaque candidat a été mesurée dans des jeux de données publiques de RNA-Seq sur la banque de donnée SRA : les banques SRR7281266, SRR7281267, SRR7281272, SRR7281273, SRR7281278, SRR7281279, SRR7281282 et SRR7281283, et exprimée en parts par million de *reads* s'alignant sur le génome (« ppm »). Certains gènes peuvent présenter plusieurs sites de complémentarité à un miARN, de niveau de conservation différents : ils apparaissent alors plusieurs fois (e.g., Matr3 / miR-320-3p). Barre d'erreur : erreur-type parmi les 8 réplicats.

constructions plasmidiques où le gène humain *Nnat* a été cloné dans son intégralité (introns compris, et avec 1592 bp de séquence en amont du site d'initiation de la transcription [TSS, *transcription*

start site] le plus précoce annoté pour *Nnat*). Nous sommes en train d'introduire, dans une région peu conservée d'un intron de cette construction, trois shRNA dirigés contre une région peu conservée de la 3' UTR de *Nnat*. Lorsque nous aurons aussi muté les sites de reconnaissance de ces shRNA dans notre plasmide, nous aurons à disposition un outil pour explorer l'influence de chaque paire de bases sur l'activité TDMD de *Nnat* dans des cellules humaines en culture.

2.6 Déterminants génomiques et épigénétiques des origines de réplication au cours du développement embryonnaire de *Cænorhabditis elegans*

Dans un projet collaboratif avec le laboratoire de Marcel Méchali, nous avons caractérisé les origines de réplication avant et après la gastrulation de l'embryon de *Cænorhabditis elegans*. Par des analyses à haut débit, nous avons mis en évidence les marques épigénétiques et les motifs de séquence associés à ces deux types d'origines de réplication, et comparé leur répartition génomique à celle des petits ARN non-codants et des gènes codants.

Nos travaux ont montré que les origines pré-gastrulation sont essentiellement enrichies dans le voisinage de promoteurs de gènes fortement transcrits. Après gastrulation, de nouvelles origines sont spécifiées par la mise en place d'un nouveau programme transcriptionnel : elles tendent à se localiser dans le voisinage d'*enhancers*, au moment justement où des types cellulaires différenciés se mettent en place, et activent, chacun, des *enhancers* variés.

Ce travail collaboratif (nous sommes trois auteurs correspondants) a été publié en 2017, et nous ne travaillons plus dessus [Rodríguez-Martínez *et al.*, 2017].

2.7 Collaborations

Outre nos projets de recherche propres, nous avons été impliqués dans un grand nombre de collaborations nationales et internationales (sollicités, le plus souvent, pour notre expertise sur les petits ARN régulateurs) :

- L'identification des miARN d'un Eucaryote basal, l'Algue brune *Ectocarpus siliculosus* (des études précédentes avaient été abusées par un artefact expérimental, et nous avons corrigé le répertoire officiel de ces miARN) [Tarver *et al.*, 2015]. **Rôle** : détection de miARN par *Northern blot* (travail partagé avec Natalia Pinzón, post-doctorante dans l'équipe). **Partenaires** : équipe de Mark Cock (UMR 8227).
- Le contrôle épigénétique de l'expression des miARN hétérosomaux au cours de la méiose (les gènes hétérosomaux sont réprimés transcriptionnellement au stade pachytène par un mécanisme épigénétique nommé *meiotic sex chromosome inactivation*; il avait semblé que les miARN échappaient à ce phénomène, mais nous avons montré qu'il s'agissait d'une mésinterprétation due à la grande stabilité post-transcriptionnelle des miARN : les miARN observés au stade pachytène sont en réalité hérités des stades précédents) [Royo *et al.*, 2015]. **Rôle** : analyse de données de microarray, et analyses statistiques sur des quantifications de RT-qPCR et d'histologie. **Partenaires** : équipe de James Turner (the Francis Crick institute, Londres).
- L'identification d'un variant naturel du gène *SHOX2* chez des patients humains atteints de fibrillation atriale, et qui fait apparaître dans ce gène un site de fixation pour un miARN, agent causal possible de la maladie [Hoffmann *et al.*, 2016]. **Rôle** : analyse de séquence pour la détection de complémentarités à la graine des miARN connus. **Partenaires** : équipe de Gudrun Rappold (université d'Heidelberg).
- La réflexion pour le développement de pratiques rigoureuses dans l'utilisation des ARN non-codants comme biomarqueurs de toxicités chimiques [Aigner *et al.*, 2016, Buesen *et al.*, 2017]. **Rôle** : expertise sur les méthodes de quantification des miARN, et leurs limites. **Partenaires** : le consortium d'experts réunis par l'ECETOC.

- L'identification des sites d'interaction d'une protéine Argonaute et de l'ARN viral HIV-1, par analyse à haut débit [Eckenfelder *et al.*, 2017]. **Rôle** : analyse informatique des données de RIP-Seq (travail partagé avec Natalia Pinzón, post-doctorante dans l'équipe). **Partenaires** : équipe de Stéphane Émiliani (institut Cochin).

La liste complète de nos collaborations est disponible sur <http://www.igh.cnrs.fr/equip/Seitz/equipe-collaborations.html>.

2.8 Place de ma recherche dans mon unité

Mon équipe (« Impact systémique des petits ARN régulateurs ») avait été créée à l'IGH en octobre 2011 en tant qu'« équipe junior », qui serait évaluée par un jury international après 5 ans d'activité. L'évaluation s'est déroulée fin 2016 et début 2017 (soumission de mon rapport écrit le 7 octobre 2016, interview par le jury le 20 janvier 2017), et mon équipe a été titularisée en « équipe senior », permanente.

L'IGH est organisé en trois départements, mon équipe appartient au département « Génétique et développement ». Nous prenons une part active à la vie scientifique du département (participation aux retraites de département tous les deux ans, invitation de conférenciers extérieurs pour des séminaires), et de l'IGH en général (participation aux retraites de l'IGH tous les deux ans, séminaires internes d'institut). Notre thématique (les fonctions biologiques des petits ARN régulateurs) est très proche de celles de deux autres équipes de l'institut (l'équipe de Séverine Chambeyron et celle de Kazufumi Mochizuki), avec lesquelles nous organisons des réunions tri-équipes un mardi sur deux, pendant 1h30.

2.9 Mobilités

Aucune au cours des 10 derniers semestres, mis à part mobilité fonctionnelle due à la promotion de mon équipe « junior » en équipe « senior », début 2017.

2.10 Distinctions scientifiques

- Prix Maurice Nicloux de la SFBBM (2018).
- Promu DR2 par le CNRS (octobre 2019).
- Prime d'encadrement doctoral et de recherche du CNRS (novembre 2019).

2.11 Publications scientifiques

(numérotées selon l'ordre d'importance que je leur accorde ; par importance décroissante)

2.11.1 Revues à comité de lecture

1. Pinzón *et al.* (2017) *Genome Res.* 27(2) : 234 (auteur correspondant).
2. Pinzón *et al.* (2019) *PLoS Genet.* 15(2) : e1007915 (auteur correspondant).
3. Rodríguez-Martínez *et al.* (2017) *Nat. Struct. Mol. Biol.* 24(3) : 290 (co-auteur correspondant).
4. Marlétaz *et al.* (2018) *Nature* 564(7734) : 64.
5. Tarver *et al.* (2015) *Nucleic Acids Res.*, 43(13) : 6384–6398.
6. Eckenfelder *et al.* (2017) *Nucleic Acids Res.* 45(7) : 4158.
7. Royo *et al.* (2015) *PLoS Genet.*, 11(10) : e1005461.
8. Hoffmann *et al.* (2016) *Basic Res. Cardiol.*, 111(3) : 36.

9. Seitz (2017) *RNA Biol.* 14(7) : 831 (auteur correspondant).
10. Seitz (2019) *Mol. Biol. Evol.* 36(7) : 1596 (auteur correspondant).
11. Amar and Seitz (2019) *Front. Genet.* 10 : 147 (co-auteur correspondant).
12. Aigner *et al.* (2016) *Regul. Toxicol. Pharmacol.*, 82 : 127–139.
13. Buesen *et al.* (2017) *Regul. Toxicol. Pharmacol.* 91 Suppl. 1 : S3.

2.11.2 Conférences invitées dans des congrès

1. 8 février 2018 à l'*International course on non-coding genome* de l'institut Curie à Paris.
2. 27 juin 2016 au *Tōkyō ncRNA symposium "Clues to non-coding RNA taxonomy"* à Tōkyō (Japon).
3. 7 novembre 2018 au congrès *SifrARN* à Nancy.
4. 12 février 2016 au *SMPGD (statistical methods for post-genomic data) annual workshop* à Lille (**keynote speaker**).
5. 4 juin 2015 à l'*ESPE science school "Non-coding RNAs in paediatric endocrinology"* à Annecy.
6. 21 novembre 2018 au 14ème congrès annuel du Cancéropôle Grand Sud-Ouest à la Grande-Motte.
7. 3 mars 2016 à l'*ECETOC workshop "ncRNAs and risk assessment science"* à Málaga (Espagne).
8. 19 octobre 2018 au colloque *Artificial intelligence in biology and health* à Montpellier.
9. 16 mai 2017 au colloque *Colloque inter-LabEx CRISPR-Cas9* à Montpellier.

2.11.3 Actes de colloques à comité de lecture

1. Présentation de 30 minutes au colloque *Quelles limites pour les technosciences en santé ?* de l'école de droit de l'université Clermont Auvergne, le 13 mars 2018 à Clermont-Ferrand (actes publiés dans le n°15 de la Revue du Centre Michel de l'Hospital, en septembre 2018).

2.11.4 Publications dans des revues sans comité

(aucune)

2.11.5 Communications à des congrès, symposium

(aucune ; le budget de mon équipe pourrait difficilement financer ma participation à un congrès auquel je ne serais pas invité par les organisateurs)

2.11.6 Séminaires, *workshops*

1. Du 6 au 10 octobre 2019, dans le cadre de l'école thématique du CNRS « InteRNAAt », dont j'étais le porteur : une dizaine d'heures de cours magistraux et deux heures de travaux pratiques sur ordinateur.
2. 4 janvier 2016 à l'École normale supérieure, à Paris, sur invitation d'Éric Meyer.
3. 8 avril 2019 à Banyuls, sur invitation d'Hector Escrivá.
4. 1^{er} décembre 2015 à Strasbourg, sur invitation de Stéphanie Blandin.
5. 26 novembre 2018 à Bordeaux, sur invitation de François Moreau-Gaudry.
6. 24 novembre 2016 à Marseille, sur invitation d'Assia Benabdallah.

2.11.7 Livres et ouvrages

(aucun)

2.11.8 Chapitres d'ouvrage

1. Mockly and Seitz (2019) *Methods Mol. Biol.* 170 : 291 (auteur correspondant).

2.11.9 Logiciels

(aucun)

2.11.10 Autres

Alors qu'en théorie, les auteurs de publications scientifiques sont censés mettre à disposition les données et les méthodes sur lesquelles ils basent leurs conclusions, en pratique il est souvent difficile de reproduire exactement les analyses publiées. Il est pourtant vital de garantir l'accès public aux détails des publications scientifiques (mise à disposition des programmes informatiques utilisés, des données brutes et des données intermédiaires), pour à la fois mettre un terme à la « crise de reproductibilité » qui fragilise le progrès scientifique, et pour permettre au reste de la communauté d'aider les auteurs à identifier et corriger leurs éventuelles erreurs (qui polluent inévitablement toute production humaine). À l'échelle de mon équipe, je mets en œuvre une politique facilitant la reproductibilité intégrale de nos résultats. Nos données brutes et intermédiaires, ainsi que les scripts que nous développons, sont accessibles sur le site web de notre équipe, et des liens en sont explicitement donnés dans nos articles. Nos scripts informatiques sont également accessibles sur la plateforme publique GitHub, dédiée au partage des programmes informatiques, et dont les adresses URL sont pérennes.

Nous mettons en ligne, sur le site web de l'équipe, les diaporamas de nos conférences, pour que les auditeurs y aient accès sans restriction après la conférence.

3 Enseignement, formation et diffusion de la culture scientifique

3.1 Thèses dirigées

Une étudiante : Sophie Mockly (depuis le 1^{er} octobre 2017).

3.1.1 Sophie Mockly

Sophie mène deux projets de recherche :

- l'évaluation de l'action de miR-34 sur la prolifération dans les cellules de Mammifères (voir section 2.3, page 14) ;
- l'identification d'ARN inducteurs de la dégradation de miARN par TDMD (voir section 2.5, page 20).

Part prise dans l'encadrement : je suis le directeur de thèse de Sophie, c'est moi qui lui enseigne les techniques qu'elle ne connaissait pas encore, en biologie moléculaire (*Northern blot*, capture sur billes magnétiques, 5' RACE, ...) et en bio-informatique (notions de phylogénie, analyses de données de séquençage à haut débit, analyses de conservation, ...).

Contenu et portée des travaux : Sophie a mis en évidence la mésinterprétation du rôle de miR-34 sur la prolifération dans les cellules humaines HCT-116 (voir section 2.3, page 14), et elle a découvert plusieurs candidats inducteurs de TDMD dans les neurones corticaux de Souris (voir section 2.5, page 20). Le programme informatique qu'elle a écrit à cette fin sera mis à la disposition du public à la publication de ses résultats.

3.2 Post-doctorants encadrés

Deux post-doctorants : Natalia Pinzón (de mars 2012 à septembre 2016) et Blaise Li (de novembre 2013 à juin 2016).

3.2.1 Natalia Pinzón

Le projet principal de Natalia a consisté en une mesure expérimentale de la compétition entre ARNm pour la fixation des miARN, par quantification absolue des miARN et des ARNm dans des cellules de Souris à différents stades de différenciation ; puis en la mutation du site d'interaction à miR-1 chez l'un des candidats ainsi identifiés (le gène *Tmsb4x*), par édition du génome, et à la mesure de l'effet de cette mutation sur l'activité de miR-1. Son second projet a consisté en la préparation de banques de Small RNA-Seq à partir de différents stades développementaux de *Branchiostoma lanceolatum*, en utilisant de protocoles permettant la détection et la distinction de petits ARN portant diverses fonctions chimiques sur leurs extrémités 5' et 3' (voir section 2.4), afin de rechercher les traces d'une activité des polymérase à ARN dépendant de l'ARN dans l'amplification du RNAi chez cette espèce.

Natalia a également été impliquée dans plusieurs collaborations (détection de miARN par *Northern blot* chez *Ectocarpus siliculosus* ; analyse des données de séquençage à haut débit d'immunoprécipitations pour déterminer la couverture de l'ARN HIV-1 en protéines Argonaute ; analyse de données génomiques pour la caractérisation des origines de répliation aux stades pré- et post-gastrula de l'embryogenèse du Nématode).

J'étais l'encadrant des travaux de Natalia, c'est moi qui lui enseignait les techniques qu'elle ne connaissait pas encore.

Contenu et portée des travaux : Natalia a pris la plus grande part dans la mise en évidence de l'effet titrateur d'ARNm abondants sur l'activité des miARN, elle est première auteur sur notre article qui décrit les faux positifs dans les prédictions de cibles de miARN [Pinzón *et al.*, 2017]. Elle a réalisé la totalité du travail de biologie moléculaire dans l'article où nous avons montré que les polymérase à ARN dépendant de l'ARN ne participent pas au RNAi chez *B. lanceolatum*, elle est également première auteur de ce travail [Pinzón *et al.*, 2019].

Devenir : Natalia est actuellement en train de réaliser un deuxième stage post-doctoral, à l'Institut Gustave Roussy (groupe de Luc Friboulet).

3.2.2 Blaise Li

Le projet principal de Blaise a consisté en une évaluation du nombre de sites de fixation de miARN plus conservés que le miARN eux-mêmes, par une recherche de complémentarité de séquence à la graine des miARN connus, et une comparaison de leur conservation phylogénétique (à partir des alignements multiples pan-génomiques de l'*UCSC Genome Browser*) à celle des miARN eux-mêmes.

Blaise était co-encadré par moi-même, et par ma collègue Séverine Chambeyron (qui dirige une autre équipe de recherche dans mon unité), et son travail s'est partagé à parts égales entre nos deux équipes. C'est moi qui ai enseigné à Blaise la bio-informatique (il savait déjà programmer, et avait une formation de biologiste ; pour son travail dans nos deux équipes, il lui restait à apprendre à analyser les données de séquençage à haut débit par des outils dédiés, et à se familiariser avec le domaine scientifique des petits ARN).

Contenu et portée des travaux : Blaise a utilisé une méthode originale (la mesure de la fréquence de « sur-conservation » des sites de fixation de miARN : lorsque le site de fixation est plus conservé que le miARN lui-même) pour démontrer que les programmes informatiques de prédiction de cibles de miARN sont abusés par des événements de conservation phylogénétique indépendants des miARN. Cette démonstration a eu un grand retentissement dans la communauté, c'était la première mise en évidence explicite de faux positifs dans les approches basées sur la génomique comparative.

Devenir : Blaise occupe maintenant un poste permanent d'ingénieur de recherche à l'Institut Pasteur.

3.3 Participations à l'enseignement

Cours magistraux sur les petits ARN régulateurs eucaryotiques (miARN, siARN et piARN) :

- 2015 : cours de 6h aux étudiants de L3 de l'École normale supérieure (titre : « Petits ARN régulateurs chez les Animaux »);
- 2016 : cours de 6h aux étudiants de L3 de l'École normale supérieure (titre : « Petits ARN régulateurs chez les Animaux »);
- 2016 : cours de 2h aux étudiants de M1 de l'École normale supérieure (titre : « *miRNA in animals* »);
- 2017 : cours de 6h aux étudiants de L3 de l'École normale supérieure (titre : « Petits ARN régulateurs chez les Animaux »);
- 2017 : cours de 2h aux étudiants de M1 de l'École normale supérieure (titre : « *miRNA in animals* »);
- 2018 : cours de 6h aux étudiants de L3 de l'École normale supérieure (titre : « Petits ARN régulateurs chez les Animaux »);
- 2019 : cours de 2h aux étudiants de M1 de l'École normale supérieure (titre : « Conséquences quantitatives de la régulation génique à l'échelle des systèmes biologiques : l'exemple des cibles de microARN »).

Intervention sur le métier de chercheur :

- 2018 : cours de 30 min aux étudiants de master et thèse de Montpellier (titre : « *A life dedicated to science : why ? how ?* »).

3.4 Participation à l'organisation de conférences

J'étais le porteur de l'école thématique du CNRS « InteRNAt », organisée à Sète du 6 au 10 octobre 2019 (35 participants).

3.5 Participation à la diffusion de l'information scientifique et technique vers des publics non spécialistes

N.B. : Les diaporamas de toutes mes interventions publiques sont en accès libre sur le site web de mon labo, leur URL est donnée en début de conférence pour que l'assistance puisse relire les diapositives à tête reposée.

Interventions orales :

- le 4 mars 2015 pour l'association « Rando, loisirs, culture » de Grabels : conférence d'1h30 (titre : « Qu'est-ce que la génétique ? »).
- le 13 janvier 2016 pour l'association « Rando, loisirs, culture » de Grabels : conférence d'1h30 (titre : « Comment l'ADN détermine le vivant »).
- le 8 octobre 2015 pour le public de la Fête de la Science à Montpellier : conférence d'1h (titre : « Qu'est-ce que la génétique ? »).
- le 24 mai 2016 pour le club « sciences » du collège Clémence Royer de Montpellier : conférence d'1h (titre : « Utilisation de la génétique de la Drosophile pour explorer les mécanismes biologiques »).
- le 25 octobre 2017 pour l'association « Rando, loisirs, culture » de Grabels : conférence d'1h30 (titre : « Sur la notion de race. Le point de vue du biologiste. »).
- le 12 octobre 2017 pour le public de la Fête de la Science à Montpellier : conférence d'1h (titre : « CRISPR/Cas9 ou les nouveaux ciseaux moléculaires »).

- le 11 janvier 2018 pour une classe de terminale du lycée St-Joseph d'Avignon : conférence de 30 min (titre : « La recherche en génétique moléculaire »).
- le 13 mars 2018 pour le public du colloque *Quelles limites pour les technosciences en santé ?* de l'école de droit de Clermont-Ferrand : conférence de 30 min (titre : « La technologie CRISPR/Cas9 »).
- le 25 avril 2018 pour le Café des sciences d'Avignon : conférence d'1h (titre : « CRISPR-Cas9 – Les ciseaux révolutionnaires de l'ADN : Dr Jekyll ou Mr Hyde ? »).
- les 6 et 7 octobre 2018 pour le public de la Fête de la Science à Montpellier : conférence d'1h chaque jour (titre : « Qu'est-ce que la génétique ? »).
- les 6 et 7 octobre 2018 pour le public de la Fête de la Science à Montpellier : conférence d'1h chaque jour (titre : « Comment l'ADN détermine le vivant »).
- les 6 et 7 octobre 2018 pour le public de la Fête de la Science à Montpellier : conférence d'1h chaque jour (titre : « Sur la notion de race. Le point de vue du biologiste. »).
- le 28 novembre 2018 pour l'association « Déclics », devant des élèves de terminale du lycée Jean Jaurès de St-Clément-de-Rivière : conférence de 30 min (titre : « Nos travaux de recherche en génétique moléculaire »).
- le 16 janvier 2019 pour le public de l'université du tiers temps à Montpellier (public essentiellement retraité) : conférence d'1h (titre : « Comment l'ADN détermine le vivant »).
- le 23 janvier 2019 pour le public de l'université du tiers temps à Montpellier (public essentiellement retraité) : conférence d'1h (titre : « Sur la notion de race. Le point de vue du biologiste. »).
- le 15 mai 2019 pour l'association « Rando, loisirs, culture » de Grabels : conférence d'1h30 (titre : « Éthique et génétique »).
- le 17 octobre 2019 pour le public du Visiatome de Marcoule : conférence d'1h (titre : « CRISPR/Cas9 : de nouveaux ciseaux moléculaires vraiment révolutionnaires ? »).
- le 28 novembre 2019 pour le public du colloque *Santé et droit de choisir* de l'école de droit de Clermont-Ferrand : conférence de 30 min (titre : « Le génome, une "donnée personnelle" d'un nouveau genre : nouvelles questions posées par la généralisation du séquençage de l'ADN »).

Interview audio, diffusée sur le podcast « Tic-talk » du Laboratoire d'informatique de Grenoble le 12 décembre 2019 :

Parallèle entre l'informatique et la génétique : sur l'édition génétique.

3.6 Interventions dans la presse écrite et audiovisuelle

Interventions écrites : *Le Monde* (1 intervention en 2015), <http://www.atlantico.fr> (1 contribution en 2015, 4 contributions en 2016, 2 contributions en 2018, 1 en 2019). Correction d'une note publiée par l'ANSES dans son Bulletin de veille scientifique n°30 (octobre 2016), rédaction de 4 notes pour les Bulletins de veille scientifique n°33, 34, 35 et 36 (parues en 2018 et à paraître en 2020).

Interview pour la radio *Aviva* le 24 octobre 2017.

3.7 Participation à des travaux d'expertise

Rapporteur pour des journaux dans le domaine de l'ARN, de la bio-informatique et de la génétique moléculaire (*Current Biology*, *eLife*, *EMBO Journal*, *EMBO Reports*, *Genome Research*, *Nucleic Acids Research*, ...) (la plupart de mes activités de *reviewing* depuis 2015 sont recensées sur <https://publons.com/researcher/1288691/herve-seitz/>).

Éditeur invité pour un numéro spécial de *Frontiers in Genetics* sur les microARN. Membre du *grant panel* « *Governance Leadership Team* » du *Cefic LRI* pour les domaines « 'omics, epigenetics and non-coding RNAs », *chairman* du comité de sélection pour l'appel d'offre "C6 - Omics and Read across" du *Cefic LRI* (2017).

Rapporteur pour des agences de financement nationales et internationales :

- KU Leuven Research Council (2019),
- ERC Starting grant (2017),
- Swiss national science program (2017 et 2018);
- Cefic Long range initiative (2017);
- ANSES (2017),
- Human Frontier Science Program (2016),
- l'ANR (2012 : panel SVSE2; 2014 : panel CE10; 2016 : panel CE12),
- l'IUF (2015),
- l'Idex Sorbonne universités (2014),
- la FRM Alsace (2011),
- le GIS IBiSA (2010).

Membre du comité de sélection des projets de la plateforme ABIC de Montpellier (2016 – 2017).

4 Transfert technologique, relations industrielles et valorisation

4.1 Contrats de recherche

En tant que partenaire :

CEFIC LRI C5 (projet « XomeTox », coordinateur : Jörg Hackermüller) (2017). Montant total : 805 337,10 €; montant alloué à mon équipe : 56 230 €. Le projet consiste à rechercher des marqueurs de toxicité (longs ARN ou petits ARN) dans le foie et la thyroïde de rats traités avec une substance toxique; l'objectif est d'améliorer les procédures de diagnostic de toxicité, en réduisant les sacrifices animaux requis, ainsi que la durée et le coût des expériences. Mon rôle est de quantifier les petits ARN dans les échantillons de foie et de thyroïde.

En tant que coordinateur :

- « Émergence » du Cancéropôle Grand Sud-Ouest (2016). Montant : 20 000 €. Le projet a consisté à caractériser l'effet anti-prolifératif de miR-34.
- Projet ARC (2019). Montant : 50 000 €. Le projet consiste à identifier les cibles de miR-21 responsables de son action sur la prolifération.

4.2 Activités de consultance

Activité de consultance pour Medesis Pharma S.A. : de mars à mai 2012, puis de mars à mai 2019 (environ une vingtaine d'heures chaque fois).

5 Encadrement, animation et management de la recherche

5.1 Direction d'équipe

Je dirige depuis octobre 2011 une équipe de recherche à l'IGH (d'abord UPR 1142, désormais UMR 9002), à Montpellier. C'était une « équipe junior », dont l'activité a été évaluée après 5 ans (fin 2016 et début 2017), et qui a été titularisée en « équipe senior », permanente (voir rubrique 2.8, en page 23). Les effectifs de l'équipe ont oscillé entre 3 et 6 personnes au cours des 10 derniers semestres.

J'accorde une très grande importance à la formation des personnes qui passent dans l'équipe. Tous les membres actuels et anciens, qui ont passé au moins un an dans l'équipe, travaillent encore dans le domaine de la recherche et de l'enseignement supérieur : L. Martinez (préparatrice à l'université

de Nantes), A. Sergeeva (scientifique au *Hematology Research Center* de Moscou), N. Pinzón (post-doctorante à l'institut Gustave Roussy) et B. Li (poste permanent d'ingénieur de recherche à l'institut Pasteur). Trois de ces quatre personnes (L. Martinez, A. Sergeeva et B. Li) bénéficient de postes permanents. D'autre part, chacun des membres actuels et anciens, qui ont passé au moins un an dans l'équipe, a publié les travaux qu'il a réalisés avec nous.

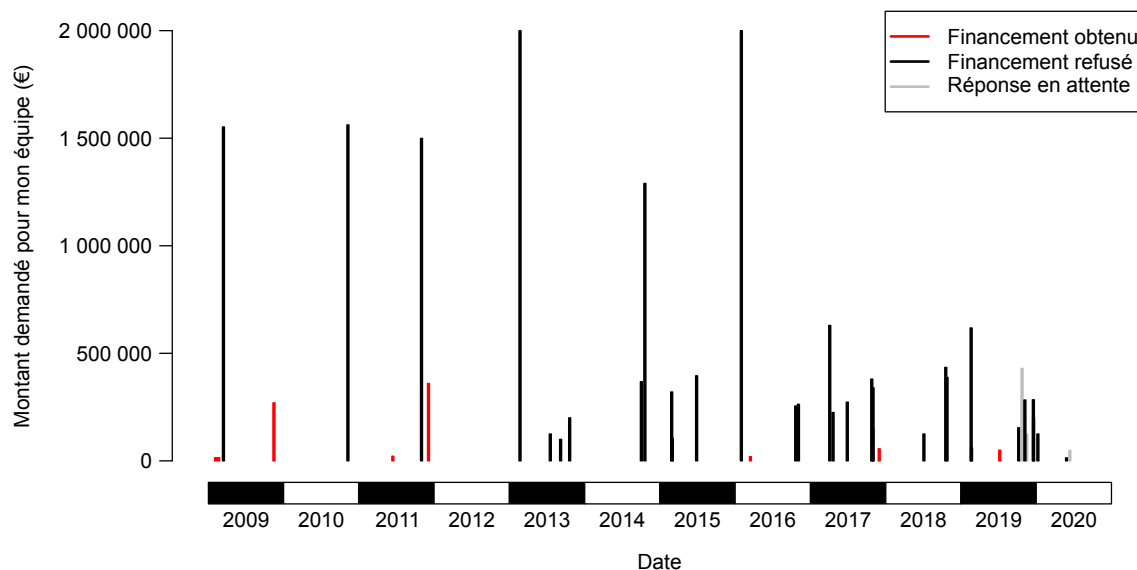


FIGURE 8 – **Historique de mes demandes de financement.** Une bonne part de mon temps de travail est consacré à ... rédiger des demandes de financement dans l'espoir (le plus souvent déçu) de réaliser les projets pour lesquels, par ailleurs, je suis salarié.

5.2 Autres responsabilités ou activités collectives au sein du CNRS

J'ai organisé, en tant que porteur, l'école thématique du CNRS « InterNAt » sur les petits ARN régulateurs (organisée du 6 au 10 octobre 2019 à Sète) : constitution du comité d'organisation et comité scientifique, choix du programme de l'école, choix et invitation des intervenants extérieurs, coordination du contenu des interventions, gestion des aspects pratiques (choix et réservation du lieu d'hébergement, ...).

6 Références

- [Aigner *et al.*, 2016] Aigner, A., Buesen, R., Gant, T., Gooderham, N., Greim, H., Hackermüller, J., Hubesch, B., Laffont, M., Marczylo, E., Meister, G., Petrick, J. S., Rasoulpour, R. J., Sauer, U. G., Schmidt, K., Seitz, H., Slack, F., Sukata, T., van der Vies, S. M., Verhaert, J., Witwer, K. W., et Poole, A. (2016). Advancing the use of noncoding RNA in regulatory toxicology : Report of an ECETOC workshop. *Regul Toxicol Pharmacol*, 82 : 127–139.
- [Allshire, 2002] Allshire, R. (2002). Molecular biology. RNAi and heterochromatin—a hushed-up affair. *Science*, 297(5588) : 1818–1819.
- [Ameres *et al.*, 2010] Ameres, S. L., Horwich, M. D., Hung, J. H., Xu, J., Ghildiyal, M., Weng, Z., et Zamore, P. D. (2010). Target RNA-directed trimming and tailing of small silencing RNAs. *Science*, 328(5985) : 1534–1539.
- [Baek *et al.*, 2008] Baek, D., Villén, J., Shin, C., Camargo, F. D., Gygi, S. P., et Bartel, D. P. (2008). The impact of microRNAs on protein output. *Nature*, 455(7209) : 64–71.
- [Bartel, 2009] Bartel, D. P. (2009). MicroRNAs : target recognition and regulatory functions. *Cell*, 136(2) : 215–233.
- [Bartel et Chen, 2004] Bartel, D. P. et Chen, C. Z. (2004). Micromanagers of gene expression : the potentially widespread influence of metazoan microRNAs. *Nat Rev Genet*, 5(5) : 396–400.
- [Bassett *et al.*, 2014] Bassett, A. R., Azzam, G., Wheatley, L., Tibbit, C., Rajakumar, T., McGowan, S., Stanger, N., Ewels, P. A., Taylor, S., Ponting, C. P., Liu, J. L., Sauka-Spengler, T., et Fulga, T. A. (2014). Understanding functional miRNA-target interactions in vivo by site-specific genome engineering. *Nat Commun*, 5 : 4640.
- [Becam *et al.*, 2011] Becam, I., Rafel, N., Hong, X., Cohen, S. M., et Milán, M. (2011). Notch-mediated repression of bantam miRNA contributes to boundary formation in the *Drosophila* wing. *Development*, 138(17) : 3781–3789.
- [Bitetti *et al.*, 2018] Bitetti, A., Mallory, A. C., Golini, E., Carrieri, C., Carreño Gutiérrez, H., Perlas, E., Pérez-Rico, Y. A., Tocchini-Valentini, G. P., Enright, A. J., Norton, W. H. J., Mandillo, S., O’Carroll, D., et Shkumatava, A. (2018). MicroRNA degradation by a conserved target RNA regulates animal behavior. *Nat Struct Mol Biol*, 25(3) : 244–251.
- [Bosson *et al.*, 2014] Bosson, A. D., Zamudio, J. R., et Sharp, P. A. (2014). Endogenous miRNA and target concentrations determine susceptibility to potential ceRNA competition. *Mol Cell*, 56(3) : 347–359.
- [Boulan *et al.*, 2013] Boulan, L., Martín, D., et Milán, M. (2013). bantam miRNA promotes systemic growth by connecting insulin signaling and ecdysone production. *Curr Biol*, 23(6) : 473–478.
- [Brennecke *et al.*, 2003] Brennecke, J., Hipfner, D. R., Stark, A., Russell, R. B., et Cohen, S. M. (2003). bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene hid in *Drosophila*. *Cell*, 113(1) : 25–36.
- [Broderick et Zamore, 2014] Broderick, J. A. et Zamore, P. D. (2014). Competitive endogenous RNAs cannot alter microRNA function in vivo. *Mol Cell*, 54(5) : 711–713.
- [Buesen *et al.*, 2017] Buesen, R., Chorley, B. N., da Silva Lima, B., Daston, G., Deferme, L., Ebbels, T., Gant, T. W., Goetz, A., Grealley, J., Gribaldo, L., Hackermüller, J., Hubesch, B., Jennen, D., Johnson, K., Kanno, J., Kauffmann, H. M., Laffont, M., McMullen, P., Meehan, R., Pemberton, M., Perdichizzi, S., Piersma, A. H., Sauer, U. G., Schmidt, K., Seitz, H., Sumida, K., Tollefsen, K. E., Tong, W., Tralau, T., van Ravenzwaay, B., Weber, R. J. M., Worth, A., Yauk, C., et Poole, A. (2017). Applying ’omics technologies in chemicals risk assessment : Report of an ECETOC workshop. *Regul Toxicol Pharmacol*, 91(Suppl 1) : S3–S13.
- [Cazalla *et al.*, 2010] Cazalla, D., Yario, T., et Steitz, J. A. (2010). Down-regulation of a host microRNA by a Herpesvirus saimiri noncoding RNA. *Science*, 328(5985) : 1563–1566.

- [Chang *et al.*, 2007] Chang, T. C., Wentzel, E. A., Kent, O. A., Ramachandran, K., Mullendore, M., Lee, K. H., Feldmann, G., Yamakuchi, M., Ferlito, M., Lowenstein, C. J., Arking, D. E., Beer, M. A., Maitra, A., et Mendell, J. T. (2007). Transactivation of miR-34a by p53 broadly influences gene expression and promotes apoptosis. *Mol Cell*, 26(5) : 745–752.
- [Cogoni et Macino, 1999] Cogoni, C. et Macino, G. (1999). Gene silencing in *Neurospora crassa* requires a protein homologous to RNA-dependent RNA polymerase. *Nature*, 399(6732) : 166–169.
- [Dalmay *et al.*, 2000] Dalmay, T., Hamilton, A., Rudd, S., Angell, S., et Baulcombe, D. C. (2000). An RNA-dependent RNA polymerase gene in *Arabidopsis* is required for posttranscriptional gene silencing mediated by a transgene but not by a virus. *Cell*, 101(5) : 543–553.
- [Dang *et al.*, 2008] Dang, V. T., Kassahn, K. S., Marcos, A. E., et Ragan, M. A. (2008). Identification of human haploinsufficient genes and their genomic proximity to segmental duplications. *Eur J Hum Genet*, 16(11) : 1350–1357.
- [de la Mata *et al.*, 2015] de la Mata, M., Gaidatzis, D., Vitanescu, M., Stadler, M. B., Wentzel, C., Scheiffele, P., Filipowicz, W., et Großhans, H. (2015). Potent degradation of neuronal miRNAs induced by highly complementary targets. *EMBO Rep*, 16(4) : 500–511.
- [Denzler *et al.*, 2014] Denzler, R., Agarwal, V., Stefano, J., Bartel, D. P., et Stoffel, M. (2014). Assessing the ceRNA hypothesis with quantitative measurements of miRNA and target abundance. *Mol Cell*, 54(5) : 766–776.
- [Denzler *et al.*, 2016] Denzler, R., McGeary, S. E., Title, A. C., Agarwal, V., Bartel, D. P., et Stoffel, M. (2016). Impact of MicroRNA levels, target-site complementarity, and cooperativity on competing endogenous RNA-regulated gene expression. *Mol Cell*, 64(3) : 565–579.
- [Doench *et al.*, 2003] Doench, J. G., Petersen, C. P., et Sharp, P. A. (2003). siRNAs can function as miRNAs. *Genes Dev*, 17(4) : 438–442.
- [Ebert et Sharp, 2010] Ebert, M. S. et Sharp, P. A. (2010). Emerging roles for natural microRNA sponges. *Curr Biol*, 20(19) : R858–R861.
- [Eckenfelder *et al.*, 2017] Eckenfelder, A., Ségéral, E., Pinzón, N., Ulveling, D., Amadori, C., Charpentier, M., Nidelet, S., Concordet, J. P., Zagury, J. F., Paillart, J. C., Berlioz-Torrent, C., Seitz, H., Emiliani, S., et Gallois-Montbrun, S. (2017). Argonaute proteins regulate HIV-1 multiply spliced RNA and viral production in a Dicer independent manner. *Nucleic Acids Res*, 45(7) : 4158–4173.
- [Ecsedi *et al.*, 2015] Ecsedi, M., Rausch, M., et Großhans, H. (2015). The *let-7* microRNA directs vulval development through a single target. *Dev Cell*, 32(3) : 335–344.
- [Friedman et Burge, 2014] Friedman, R. C. et Burge, C. B. (2014). MicroRNA target finding by comparative genomics. *Methods Mol Biol*, 1097 : 457–476.
- [Friedman *et al.*, 2009] Friedman, R. C., Farh, K. K., Burge, C. B., et Bartel, D. P. (2009). Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res*, 19(1) : 92–105.
- [Fuchs Wightman *et al.*, 2018] Fuchs Wightman, F., Giono, L. E., Fededa, J. P., et de la Mata, M. (2018). Target RNAs strike back on microRNAs. *Front Genet*, 9 : 435.
- [Garcia *et al.*, 2011] Garcia, D. M., Baek, D., Shin, C., Bell, G. W., Grimson, A., et Bartel, D. P. (2011). Weak seed-pairing stability and high target-site abundance decrease the proficiency of lsy-6 and other microRNAs. *Nat Struct Mol Biol*, 18(10) : 1139–1146.
- [Ghildiyal et Zamore, 2009] Ghildiyal, M. et Zamore, P. D. (2009). Small silencing RNAs : an expanding universe. *Nat Rev Genet*, 10(2) : 94–108.
- [Goljanek-Whysall *et al.*, 2012] Goljanek-Whysall, K., Pais, H., Rathjen, T., Sweetman, D., Dalmay, T., et Münsterberg, A. (2012). Regulation of multiple target genes by miR-1 and miR-206 is pivotal for C2C12 myoblast differentiation. *J Cell Sci*, 125(Pt 15) : 3590–3600.

- [Gu *et al.*, 2009] Gu, W., Shirayama, M., Conte, D. J., Vasale, J., Batista, P. J., Claycomb, J. M., Moresco, J. J., Youngman, E. M., Keys, J., Stoltz, M. J., Chen, C. C., Chaves, D. A., Duan, S., Kasschau, K. D., Fahlgren, N., Yates, J. R., Mitani, S., Carrington, J. C., et Mello, C. C. (2009). Distinct argonaute-mediated 22G-RNA pathways direct genome surveillance in the *C. elegans* germline. *Mol Cell*, 36(2) : 231–244.
- [Guo *et al.*, 2019] Guo, Z., Li, Y., et Ding, S. W. (2019). Small RNA-based antimicrobial immunity. *Nat Rev Immunol*, 19(1) : 31–44.
- [Hammond et Keller, 2005] Hammond, T. M. et Keller, N. P. (2005). RNA silencing in *Aspergillus nidulans* is independent of RNA-dependent RNA polymerases. *Genetics*, 169(2) : 607–617.
- [Han *et al.*, 2009] Han, T., Manoharan, A. P., Harkins, T. T., Bouffard, P., Fitzpatrick, C., Chu, D. S., Thierry-Mieg, D., Thierry-Mieg, J., et Kim, J. K. (2009). 26G endo-siRNAs regulate spermatogenic and zygotic gene expression in *Caenorhabditis elegans*. *Proc Natl Acad Sci USA*, 106(44) : 18674–18679.
- [He *et al.*, 2007] He, L., He, X., Lim, L. P., de Stanchina, E., Xuan, Z., Liang, Y., Xue, W., Zender, L., Magnus, J., Ridzon, D., Jackson, A. L., Linsley, P. S., Chen, C., Lowe, S. W., Cleary, M. A., et Hannon, G. J. (2007). A microRNA component of the p53 tumour suppressor network. *Nature*, 447(7148) : 1130–1134.
- [Hipfner *et al.*, 2002] Hipfner, D. R., Weigmann, K., et Cohen, S. M. (2002). The *bantam* gene regulates *Drosophila* growth. *Genetics*, 161(4) : 1527–1537.
- [Hoffmann *et al.*, 2016] Hoffmann, S., Clauss, S., Berger, I. M., Weiß, B., Montalbano, A., Röth, R., Bucher, M., Klier, I., Wakili, R., Seitz, H., Schulze-Bahr, E., Katus, H. A., , Flachsbart, F., Nebel, A., Guenther, S. P., Bagaev, E., Rottbauer, W., Kääb, S., Just, S., et Rappold, G. (2016). Coding and non-coding variants in the SHOX2 gene in patients with early-onset atrial fibrillation. *Basic Res Cardiol*, 111(3) : 36.
- [Huang *et al.*, 2010] Huang, N., Lee, I., Marcotte, E. M., et Hurles, M. E. (2010). Characterising and predicting haploinsufficiency in the human genome. *PLoS Genet*, 6(10) : e1001154.
- [Iwakawa et Tomari, 2015] Iwakawa, H. O. et Tomari, Y. (2015). The functions of microRNAs : mRNA decay and translational repression. *Trends Cell Biol*, 25(11) : 651–665.
- [Khvorova *et al.*, 2003] Khvorova, A., Reynolds, A., et Jayasena, S. D. (2003). Functional siRNAs and miRNAs exhibit strand bias. *Cell*, 115(2) : 209–216.
- [Kingston et Bartel, 2019] Kingston, E. R. et Bartel, D. P. (2019). Global analyses of the dynamics of mammalian microRNA metabolism. *Genome Res*, 29(11) : 1777–1790.
- [Kleaveland *et al.*, 2018] Kleaveland, B., Shi, C. Y., Stefano, J., et Bartel, D. P. (2018). A network of noncoding regulatory RNAs acts in the mammalian brain. *Cell*, 174(2) : 350–362.
- [Krol *et al.*, 2010] Krol, J., Busskamp, V., Markiewicz, I., Stadler, M. B., Ribi, S., Richter, J., Duebel, J., Bicker, S., Fehling, H. J., Schübeler, D., Oertner, T. G., Schratt, G., Bibel, M., Roska, B., et Filipowicz, W. (2010). Characterizing light-regulated retinal microRNAs reveals rapid turnover as a common property of neuronal microRNAs. *Cell*, 141(4) : 618–631.
- [Leaman *et al.*, 2005] Leaman, D., Chen, P. Y., Fak, J., Yalcin, A., Pearce, M., Unnerstall, U., Marks, D. S., Sander, C., Tuschl, T., et Gaul, U. (2005). Antisense-mediated depletion reveals essential and specific functions of microRNAs in *Drosophila* development. *Cell*, 121(7) : 1097–1108.
- [Lewis *et al.*, 2018] Lewis, S. H., Quarles, K. A., Yang, Y., Tanguy, M., Frézal, L., Smith, S. A., Sharma, P. P., Cordaux, R., Gilbert, C., Giraud, I., Collins, D. H., Zamore, P. D., Miska, E. A., Sarkies, P., et Jiggins, F. M. (2018). Pan-arthropod analysis reveals somatic piRNAs as an ancestral defence against transposable elements. *Nat Ecol Evol*, 2(1) : 174–181.
- [Li *et al.*, 2018] Li, H., Bowling, A. J., Gandra, P., Rangasamy, M., Pence, H. E., McEwan, R. E., Khajuria, C., Siegfried, B. D., et Narva, K. E. (2018). Systemic RNAi in western corn rootworm, *Diabrotica virgifera virgifera*, does not involve transitive pathways. *Insect Sci*, 25(1) : 45–56.

- [Marcinowski *et al.*, 2012] Marcinowski, L., Tanguy, M., Krmpotic, A., Rädle, B., Lisnić, V. J., Tud-denham, L., Chane-Woon-Ming, B., Ruzsics, Z., Erhard, F., Benkartek, C., Babic, M., Zimmer, R., Trgovcich, J., Koszinowski, U. H., Jonjic, S., Pfeffer, S., et Dölken, L. (2012). Degradation of cellular miR-27 by a novel, highly abundant viral transcript is important for efficient virus replication *in vivo*. *PLoS Pathog*, 8(2) : e1002510.
- [Marker *et al.*, 2010] Marker, S., Le, Mouël, A., Meyer, E., et Simon, M. (2010). Distinct RNA-dependent RNA polymerases are required for RNAi triggered by double-stranded RNA versus truncated transgenes in *Paramecium tetraurelia*. *Nucleic Acids Res*, 38(12) : 4092–4107.
- [Marlétaz *et al.*, 2018] Marlétaz, F., Firbas, P. N., Maeso, I., Tena, J. J., Bogdanovic, O., Perry, M., Wyatt, C. D. R., de la Calle-Mustienes, E., Bertrand, S., Burguera, D., Acemel, R. D., van Heeringen, S. J., Naranjo, S., Herrera-Ubeda, C., Skvortsova, K., Jimenez-Gancedo, S., Aldea, D., Marquez, Y., Buono, L., Kozmikova, I., Permanyer, J., Louis, A., Albuixech-Crespo, B., Le Pétillon, Y., Leon, A., Subirana, L., Balwierz, P. J., Duckett, P. E., Farahani, E., Aury, J. M., Mangenot, S., Wincker, P., Albalat, R., Benito-Gutiérrez, E., Cañestro, C., Castro, F., D’Aniello, S., Ferrier, D. E. K., Huang, S., Laudet, V., Marais, G. A. B., Pontarotti, P., Schubert, M., Seitz, H., Somorjai, I., Takahashi, T., Mirabeau, O., Xu, A., Yu, J. K., Carninci, P., Martinez-Morales, J. R., Roest-Crollius, H., Kozmik, Z., Weirauch, M. T., Garcia-Fernández, J., Lister, R., Lenhard, B., Holland, P. W. H., Escrivà, H., Gómez-Skarmeta, J. L., et Irimia, M. (2018). Amphioxus functional genomics and the origins of vertebrate gene regulation. *Nature*, 564(7734) : 64–70.
- [Martienssen et Moazed, 2015] Martienssen, R. et Moazed, D. (2015). RNAi and heterochromatin assembly. *Cold Spring Harb Perspect Biol*, 7(8) : a019323.
- [Mittal et Zavolan, 2014] Mittal, N. et Zavolan, M. (2014). Seq and CLIP through the miRNA world. *Genome Biol*, 15(1) : 202.
- [Mockly et Seitz, 2019] Mockly, S. et Seitz, H. (2019). Inconsistencies and limitations of current microRNA target identification methods. *Methods Mol Biol*, 1970 : 291–314.
- [Motamedi *et al.*, 2004] Motamedi, M. R., Verdel, A., Colmenares, S. U., Gerber, S. A., Gygi, S. P., et Moazed, D. (2004). Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. *Cell*, 119(6) : 789–802.
- [Mourrain *et al.*, 2000] Mourrain, P., Béclin, C., Elmayan, T., Feuerbach, F., Godon, C., Morel, J.-B., Jouette, D., Lacombe, A. M., Nikic, S., Picault, N., Remoue, K., Sanial, M., Vo, T. A., et Vaucheret, H. (2000). *Arabidopsis SGS2* and *SGS3* genes are required for posttranscriptional gene silencing and natural virus resistance. *Cell*, 101(5) : 533–542.
- [O’Donnell *et al.*, 2005] O’Donnell, K. A., Wentzel, E. A., Zeller, K. I., Dang, C. V., et Mendell, J. T. (2005). c-Myc-regulated microRNAs modulate E2F1 expression. *Nature*, 435(7043) : 839–843.
- [Pak et Fire, 2007] Pak, J. et Fire, A. (2007). Distinct populations of primary and secondary effectors during RNAi in *C. elegans*. *Science*, 315(5809) : 241–244.
- [Péligsson *et al.*, 2007] Péligsson, A., Sarot, E., Payen-Groschêne, G., et Bucheton, A. (2007). A novel repeat-associated small interfering RNA-mediated silencing pathway downregulates complementary sense gypsy transcripts in somatic cells of the *Drosophila* ovary. *J Virol*, 81(4) : 1951–1960.
- [Pinzón *et al.*, 2019] Pinzón, N., Bertrand, S., Subirana, L., Busseau, I., Escrivà, H., et Seitz, H. (2019). Functional lability of RNA-dependent RNA polymerases in animals. *PLoS Genet*, 15(2) : e1007915.
- [Pinzón *et al.*, 2017] Pinzón, N., Li, B., Martinez, L., Sergeeva, A., Presumey, J., Apparailly, F., , et Seitz, H. (2017). microRNA target prediction programs predict many false positives. *Genome Res*, 27(2) : 234–245.
- [Raver-Shapira *et al.*, 2007] Raver-Shapira, N., Marciano, E., Meiri, E., Spector, Y., Rosenfeld, N., Moskovits, N., Bentwich, Z., et Oren, M. (2007). Transcriptional activation of miR-34a contributes to p53-mediated apoptosis. *Mol Cell*, 26(5) : 731–743.

- [Reichholf *et al.*, 2019] Reichholf, B., Herzog, V. A., Fasching, N., Manzenreither, R. A., Sowemimo, I., et Ameres, S. L. (2019). Time-resolved small RNA sequencing unravels the molecular principles of microRNA homeostasis. *Mol Cell*, 75(4) : 756–768.
- [Rodríguez-Martínez *et al.*, 2017] Rodríguez-Martínez, M., Pinzón, N., Ghommidh, C., Beyne, E., Seitz, H., Cayrou, C., et Méchali, M. (2017). The gastrula transition reorganizes replication-origin selection in *Caenorhabditis elegans*. *Nat Struct Mol Biol*, 24(3) : 290–299.
- [Rosenbloom *et al.*, 2015] Rosenbloom, K. R., Armstrong, J., Barber, G. P., Casper, J., Clawson, H., Diekhans, M., Dreszer, T. R., Fujita, P. A., Guruvadoo, L., Haeussler, M., Harte, R. A., Heitner, S., Hickey, G., Hinrichs, A. S., Hubley, R., Karolchik, D., Learned, K., Lee, B. T., Li, C. H., Miga, K. H., Nguyen, N., Paten, B., Raney, B. J., Smit, A. F., Speir, M. L., Zweig, A. S., Haussler, D., Kuhn, R. M., et Kent, W. J. (2015). The UCSC Genome Browser database : 2015 update. *Nucleic Acids Res*, 43(Database issue) : D670–681.
- [Royo *et al.*, 2015] Royo, H., Seitz, H., Ellnati, E., Peters, A. H., Stadler, M. B., et Turner, J. M. (2015). Silencing of X-linked microRNAs by meiotic sex chromosome inactivation. *PLoS Genet*, 11(10) : e1005461.
- [Ruby *et al.*, 2006] Ruby, J. G., Jan, C., Player, C., Axtell, M. J., Lee, W., Nusbaum, C., Ge, H., et Bartel, D. P. (2006). Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in *C. elegans*. *Cell*, 127(6) : 1193–1207.
- [Schwarz *et al.*, 2003] Schwarz, D. S., Hutvágner, G., Du, T., Xu, Z., Aronin, N., et Zamore, P. D. (2003). Asymmetry in the assembly of the RNAi enzyme complex. *Cell*, 115(2) : 199–208.
- [Seitz, 2009] Seitz, H. (2009). Redefining microRNA targets. *Curr Biol*, 19(10) : 870–873.
- [Seitz, 2017] Seitz, H. (2017). Issues in current microRNA target identification methods. *RNA Biol*, 14(7) : 831–834.
- [Selbach *et al.*, 2008] Selbach, M., Schwanhäusser, B., Thierfelder, N., Fang, Z., Khanin, R., et Rajewsky, N. (2008). Widespread changes in protein synthesis induced by microRNAs. *Nature*, 455(7209) : 58–63.
- [Sijen *et al.*, 2001] Sijen, T., Fleenor, J., Simmer, F., Thijssen, K. L., Parrish, S., Timmons, L., Plasterk, R. H., et Fire, A. (2001). On the role of RNA amplification in dsRNA-triggered gene silencing. *Cell*, 107(4) : 465–476.
- [Sijen *et al.*, 2007] Sijen, T., Steiner, F. A., Thijssen, K. L., et Plasterk, R. H. (2007). Secondary siRNAs result from unprimed RNA synthesis and form a distinct class. *Science*, 315(5809) : 244–247.
- [Smardon *et al.*, 2000] Smardon, A., Spoerke, J. M., Stacey, S. C., Klein, M. E., Mackin, N., et Maine, E. M. (2000). EGO-1 is related to RNA-directed RNA polymerase and functions in germline development and RNA interference in *C. elegans*. *Curr Biol*, 10(4) : 169–178.
- [Tarver *et al.*, 2015] Tarver, J. E., Cormier, A., Pinzón, N., Taylor, R. S., Carré, W., Strittmatter, M., Seitz, H., Coelho, S. M., et Cock, J. M. (2015). microRNAs and the evolution of complex multicellularity : identification of a large, diverse complement of microRNAs in the brown alga *Ectocarpus*. *Nucleic Acids Res*, 43(13) : 6384–6398.
- [Vasale *et al.*, 2010] Vasale, J. J., Gu, W., Thivierge, C., Batista, P. J., Claycomb, J. M., Youngman, E. M., Duchaine, T. F., Mello, C. C., et Conte, Jr., D. (2010). Sequential rounds of RNA-dependent RNA transcription drive endogenous small-RNA biogenesis in the ERGO-1/Argonaute pathway. *Proc Natl Acad Sci USA*, 107(8) : 3582–3587.
- [Viswanathan *et al.*, 2008] Viswanathan, S. R., Daley, G. Q., et Gregory, R. I. (2008). Selective blockade of microRNA processing by Lin28. *Science*, 320(5872) : 97–100.
- [Voinnet, 2008] Voinnet, O. (2008). Use, tolerance and avoidance of amplified RNA silencing by plants. *Trends Plant Sci*, 13(7) : 317–328.

[Wee *et al.*, 2012] Wee, L. M., Flores-Jasso, C. F., Salomon, W. E., et Zamore, P. D. (2012). Argonaute divides its RNA guide into domains with distinct functions and RNA-binding properties. *Cell*, 151(5) : 1055–1067.