

Rapport d'activité (janvier 2009 - décembre 2013)

Hervé Seitz

Table des matières

1	<i>Curriculum vitae</i>	3
1.1	Informations générales	3
1.2	Études et expérience post-doctorale	3
1.3	Enseignement	3
1.4	Autres activités professionnelles	4
1.5	Prix et distinctions	4
1.6	Diffusion des savoirs	4
1.7	Collaborations internationales	5
1.8	Activités non professionnelles	5
2	Recherche scientifique	6
2.1	Fonction des interactions entre microARN et ARNm-cibles à l'échelle du système	6
2.2	Histoire évolutive des petits ARN régulateurs chez les Métazoaires	13
2.3	Déterminants structuraux de la biogenèse des microARN	15
2.4	Le rôle du nucléotide 5' des microARN	15
2.5	Collaborations avec des partenaires académiques	15
2.6	Place de ma recherche dans celle de mon unité, mobilité géographique	16
2.7	Distinction scientifique	16
2.8	Publications scientifiques des 10 derniers semestres	17
2.8.1	Revue à comité de lecture	17
2.8.2	Conférences invitées dans des congrès	17
2.8.3	Actes de colloques à comité de lecture	18
2.8.4	Publications dans des revues sans comité	18
2.8.5	Communications à des congrès, symposiums	18
2.8.6	Séminaires, <i>workshops</i>	18
2.8.7	Livres et ouvrages	19
2.8.8	Chapitres d'ouvrages	19
2.8.9	Logiciels	19
2.8.10	Autres	19
3	Enseignement, formation et diffusion de la culture scientifique	20
3.1	Thèse et post-doctorats	20
3.2	Participation à l'enseignement	20
3.3	Participation à l'organisation de conférences	21
3.4	Participation à des revues ou ouvrages de vulgarisation	21
3.5	Interventions dans la presse écrite et audiovisuelle	21
3.6	Participation à des travaux d'expertise	22

4	Transfert technologique, relations industrielles et valorisation	23
5	Encadrement, animation et management de la recherche	24
6	Objectifs / Projet de recherche	25
7	Références	28

1 *Curriculum vitæ*

1.1 Informations générales

Né le 13 décembre 1977 à Angers (Maine-et-Loire).

Adresse personnelle : Apt. 107, Bastides de l'oliveraie, rue de l'oliveraie, 34790 Grabels.

Adresse professionnelle : IGH (UPR 1142 du CNRS), 141 rue de la Cardonille, 34396 Montpellier.

email : herve.seitz@igh.cnrs.fr

site web de l'équipe : <http://www.igh.cnrs.fr/FR/equipe-detail.php?id=72>

1.2 Études et expérience post-doctorale

- 1995 : Baccalauréat S, option maths, mention « très bien » (lycée H. Bergson, Angers).
- 1995-1997 : Maths sup. et spé. bio. (lycée G. Clemenceau, Nantes).
- 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).
- 2000-2001 : stages avec M. Dreyfus (caractérisation de mutants de la RNase E chez *Escherichia coli*; 3 mois), puis J. Vinh (analyses d'ARN en spectrométrie de masse; 3 mois).
- 2001-2004 : thèse (financement « allocation couplée »); titre : « Empreinte génomique parentale et petits ARN non-codants » (université Toulouse III Paul Sabatier) (directeur de thèse : Jérôme Cavaillé; 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 C.N.R.S.; 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.

1.3 Enseignement

- Moniteur en biologie moléculaire à l'Université Toulouse III Paul Sabatier (≈ 70 heures de travaux dirigés et de travaux pratiques par an, pendant trois ans; enseignement en L2, L3 et M1) : 2001-2004.
- Cours sur les petits ARN régulateurs eucaryotiques (d'une à trois interventions par an, depuis 2007, pour étudiants en M2R ou en thèse, à Toulouse, Paris, Strasbourg, Lyon et Rennes; cours de 2 à 3h chacun).
- Intervention d'une heure au *International course on Non-coding genome* de l'Institut Curie, le 11 décembre 2012; 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.

- Sessions d’initiation à l’utilisation d’Unix et à la bio-informatique (en cinq leçons) à destination du personnel de l’University of Massachusetts Medical School (printemps et été 2008), et de l’IFR 109 (Toulouse ; printemps et été 2009). Sessions d’initiation à l’utilisation d’Unix et à la bio-informatique (en deux leçons par groupe, pour 3 groupes d’une vingtaine de participants) à destination du personnel de l’IGH et des instituts du voisinage (Montpellier ; hiver 2013-2014).
- Cours « Les statistiques en biologie moléculaire », à destination du personnel de l’IFR 109 (Toulouse ; printemps et été 2010), également donné à l’University of Massachusetts Medical School en mars 2010, et à l’IGH (Montpellier) en avril 2010. Une version abrégée de ce cours est disponible en deux articles dans les numéros de juillet et octobre 2010 de *Regards sur la Biochimie* : <http://sfbbm.fr/index.php/publications/regard-sur-la-biochimie2>.
- Rapporteur pour les thèses de Caroline Jacquier (directeur de thèse : Christophe Antoniewski, Institut Pasteur, Paris ; 28 juin 2010), de Chongjian Chen (directeurs de thèse : Daniel Gautheret, université Paris Sud, Orsay, et Liang-Hu Qu, Sun Yat-Sen University, Canton, Chine ; 17 décembre 2010), d’Antoine Molaro (directeurs de thèse : Phillip Avner, Institut Pasteur, Paris, et Gregory Hannon, Cold Spring Harbor Laboratory, NY, États-Unis ; 9 mai 2012), d’Ilyass Zniber (directeur de thèse : Denis Dupuy, IECB, Bordeaux ; 17 décembre 2012), de Lue Huang (directeur de thèse : François Dautry, ENS Cachan ; 19 décembre 2012) .
- Examineur pour les thèses de Nicolas Greliche (directeur de thèse : David-Alexandre Trégouët ; 18 février 2013) et de Thomas Grentzinger (directrice de thèse : Séverine Chambeyron ; 13 juin 2013).
- Rapporteur pour l’HDR de Séverine Chambeyron (directeur des recherches : Alain Bucheton, IGH, Montpellier ; 11 mai 2011).
- Membre du jury de recrutement d’un maître de conférence pour l’université Toulouse III Paul Sabatier (poste n°64MCF0300) en avril et mai 2013.

1.4 Autres activités professionnelles

Rapporteur pour les journaux *Biochimie*, *BioEssays*, *Cell Reports*, *Cell Research*, *Current Biology*, *EMBO Reports*, *Genome Biology*, *Genome Research*, *Genomics*, *Molecular Oncology*, *Nucleic Acids Research*, *PLoS One*, *RNA*, *RNA Biology* et *Silence*.

Rapporteur pour l’appel d’offre « Deep-sequencing » du GIS IBiSA (Génoscope et Centre national de séquençage) (2010). Rapporteur pour l’appel d’offre de la FRM de la région Alsace (2011). Rapporteur pour le programme « ANR Blanche » de l’ANR, panel SVSE2 (2012).

Membre du comité de programme du congrès JOBIM (Journées Ouvertes en Biologie, Informatique et Mathématiques) 2012 et 2013.

1.5 Prix et distinctions

Prime d’excellence scientifique du CNRS (décembre 2010).

1.6 Diffusion des savoirs

Cours de deux heures, sur le thème « Petits ARN régulateurs chez les Eucaryotes » le 9 juin 2010, à destination d’enseignants en SVT dans le secondaire (formation organisée par l’IRFEC, à Toulouse).

Intervention de 30 minutes, sur les petits ARN régulateurs chez les Animaux, le 19 novembre 2010, au Carrefour Eurobiomed (journée de rencontre entre industriels et chercheurs académiques), à Montpellier.

Activités de vulgarisation : interventions sur le thème de « la cellule » auprès de scolaires (quatre groupes d’élèves de terminale, un groupe d’élèves de 6ème et de CE2), organisées par l’association « Connaissances », 31 mars et 1^{er} avril 2011.

Fête de la science 2013 : intervention de 30 minutes sur le thème « Qu'est-ce que la génétique ? » les 10 et 11 octobre 2013 (devant des groupes de scolaires) et le 12 octobre 2013 (devant le grand public).

Outre les activités de vulgarisation, je tâche d'aider mes collègues biologistes par des formations, formelles ou informelles, dans des domaines qui sont très utiles aux biologistes, mais encore mal connus (la bio-informatique et les statistiques). Je pense que, dans l'intérêt des scientifiques eux-mêmes, 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 mets en accès libre les formations que je donne à mes collègues, pour qu'elles servent au plus grand nombre :

- Un cours d'apprentissage d'Unix et de bio-informatique à destination des biologistes moléculaires (donné à mes collègues de l'University of Massachusetts Medical School au printemps et à l'été 2008 ; à mes collègues de l'IFR 109, à Toulouse, au printemps et à l'été 2009, et à mes collègues de l'IGH, à Montpellier, en décembre 2013 ; disponible sur <http://www.umassmed.edu/Content.aspx?id=43466>, sur <http://www-lbme.biotoul.fr/formations.html> et sur <http://www.igh.cnrs.fr/equip/Seitz/equipe-tutoriels.html>).
- Un cours d'apprentissage des statistiques à destination des biologistes moléculaires (donné à mes collègues de l'IFR 109 au printemps 2010, à l'University of Massachusetts Medical School en mars 2010 et à l'IGH de Montpellier en avril 2010 ; disponible sur <http://www-lbme.biotoul.fr/formations.html> et sur <http://www.igh.cnrs.fr/equip/Seitz/equipe-tutoriels.html>).

Pour faciliter la diffusion des bonnes pratiques en statistiques, j'ai également publié une version résumée du cours d'apprentissage des statistiques dans la revue *Regard sur la biochimie*, en accès libre (numéros de juillet et octobre 2010) : <http://sfbbm.fr/index.php/publications/regard-sur-la-biochimie2>.

1.7 Collaborations internationales

Depuis 2009 : avec le Prof. Y. Tomari (université de Tōkyō, Japon). Publications communes : voir Kawamata *et al.*, 2009 et Tsutsumi *et al.*, 2011, dans la liste des productions scientifiques.

Depuis 2010 : avec le Prof. U. Technau (université de Vienne, Autriche). Publication commune : en révision à *Genome Research*.

1.8 Activités non professionnelles

Ultra-marathon (médaille de bronze du championnat de France 2012 de 100 km ; troisième des 100 km de Millau 2012 ; deuxième des 100 km de Millau 2013).

2 Recherche scientifique

2.1 Fonction des interactions entre microARN et ARNm-cibles à l'échelle du système

Les microARN (« miRNA ») sont de petits (≈ 22 nt) ARN régulateurs, qui guident un complexe effecteur vers des cibles ARNm spécifiques : en s'appariant (souvent imparfaitement) à l'ARN-cible, le miRNA provoque sa répression (catalysée par les protéines associées au miRNA), par une variété de mécanismes (Chekulaeva et Filipowicz, 2009). Des programmes informatiques ont été écrits par plusieurs laboratoires, pour prédire les cibles de miRNA : la plupart de ces programmes recherchent des complémentarités imparfaites aux miRNA, et sélectionnent celles qui sont conservées dans l'évolution. De façon unanime, ces programmes prédisent des milliers de cibles (typiquement, quelques dizaines de cibles par miRNA chez la Drosophile, quelques centaines chez les Mammifères, or il existe quelques centaines de miRNA chez ces espèces) ; notamment chez les Mammifères, plus de 60% des gènes codants portent des sites de complémentarité aux miRNA, et ces sites ont été sélectivement conservés au cours de l'évolution, ce qui suggère que les miRNA régulent plus de la moitié des gènes codants de Mammifères (Friedman *et al.*, 2009).

Au début de cette période de 10 semestres (en février 2009), je soumettais un article purement théorique, sans donnée expérimentale (je venais de rentrer en France et de prendre mon poste au CNRS en janvier), pour proposer une nouvelle interprétation de la fonction biologique de l'interaction entre les miRNA et les ARNm. Trois paradoxes restaient en effet inexpliqués :

1. Dans les trois cas où des cibles de miRNA avaient pu être identifiées par la génétique, expérimentalement, il s'est avéré que tous les phénotypes du mutant du miRNA étaient dus à la dérégulation d'une unique cible. Ces expériences suggéraient donc que ces miRNA n'avaient qu'une cible chacun, alors que les prédictions informatiques en identifient des dizaines (et notamment pour les trois miRNA en question).
2. L'amplitude de la répression guidée par les miRNA est très faible (moins d'un facteur 2 en général), alors que l'activité de la plupart des gènes est robuste vis-à-vis de petites fluctuations de l'expression de ces gènes (par exemple, la plupart des gènes sont haplo-suffisants chez les Métazoaires).
3. Alors que les miRNA peuvent être très conservés chez les Animaux (notamment, entre Protostomiens et Deutérostomiens), la grande majorité de leurs cibles prédites sont, elles, différentes entre les différents phyla : pourquoi le miRNA aurait-il été conservé, si ses fonctions régulatrices ne le sont pas ?

L'article que je soumettais proposait une nouvelle interprétation : la plupart des cibles prédites informatiquement seraient, en réalité, insensibles à la (faible) répression guidée par les miRNA. Seuls quelques gènes, dont l'activité est suffisamment sensible aux fluctuations de leur expression, seraient de véritables cibles de miRNA, au sens fonctionnel (et ils seraient responsables des phénotypes observés chez les mutants de miRNA). Il fallait toutefois expliquer pourquoi l'interaction entre miRNA et cibles prédites est conservée dans l'évolution (par construction, toutes les cibles prédites portent des sites conservés, puisque c'est comme ça qu'elles sont identifiées par les programmes informatiques).

J'ai donc proposé que cette fonction consistait en la titration des miRNA (en fixant le miRNA, ces ARNm l'empêcheraient d'aller interagir avec ses quelques vraies cibles). Dans les cellules où l'inactivation du miRNA est bénéfique, le transcriptome serait donc soumis à une pression de sélection pour acquérir, et conserver, des sites d'interaction au miRNA. Ces « pseudo-cibles » ne seraient donc pas des cibles de miRNA au sens fonctionnel, mais elles porteraient quand même des sites d'interaction au miRNA, conservés au cours de l'évolution (au moins parmi les espèces proches), et seraient donc prises pour des cibles, par erreur, par les programmes de prédiction informatique. Il faut noter que cette activité de titration par les pseudo-cibles est indépendante de l'identité de l'ARNm qui porte les sites de reconnaissance du miRNA : de cette manière, sur de longues distances évolutives, la liste

de pseudo-cibles pourrait varier, les pseudo-cibles étant remplacées, dans leur rôle d'inhibiteur de miRNA, par d'autres ARNm partageant un patron d'expression similaire.

Cet article, qui proposait donc une solution unique aux trois paradoxes, tout en restant parfaitement compatible avec l'ensemble des observations expérimentales, a été publié dans *Current Biology* au printemps 2009 (voir Seitz, 2009 dans la liste de productions scientifiques). Il propose de revisiter la définition des cibles de miRNA : alors que des milliers de gènes sont effectivement affectés par les miRNA à l'échelle moléculaire (l'abondance de leurs protéines est effectivement réprimée d'un facteur 2 environ), cet effet microscopique ne se traduit en phénotype macroscopique que pour les gènes les plus sensibles (ce seraient les cibles véritables du miRNA). La différence de sensibilité des phénotypes dépendrait des mécanismes d'intégration du microscopique au macroscopique (les boucles de rétrocontrôle négatif ou positif peuvent atténuer ou amplifier la propagation de la régulation ; l'abondance relative des partenaires moléculaires d'une protéine déterminera si elle est limitante, donc susceptible d'être sensible au dosage ; *etc*) : la compréhension de ces phénomènes, qui relèvent de la nouvelle discipline de « biologie des systèmes », me semble devoir permettre de distinguer les vraies cibles de miRNA, des pseudo-cibles inhibitrices.

Mon projet de recherche essentiel consistait donc à mettre cette nouvelle hypothèse à l'épreuve, et à la confronter à la théorie précédente (qui affirmait que tous les ARNm avec une interaction conservée aux miRNA sont des cibles de miRNA, fonctionnellement affectées par le petit ARN). J'avais rejoint, en janvier 2009, l'équipe de Jérôme Cavallé, au LBME (UMR 5099 CNRS et université Paul Sabatier, à Toulouse), et j'allais développer mon projet dans le cadre de son équipe (qui était principalement intéressée par la fonction des miRNA chez les Mammifères).

Mes premières analyses consistaient à mesurer l'effet de titration des miRNA sur l'activité d'un rapporteur fluorescent, dans des cellules de *Drosophile* en culture. En réprimant, par RNAi, des cibles prédites pour le miRNA appelé *bantam*, nous voulions mesurer l'éventuelle dérégulation d'un rapporteur GFP dont la 3' UTR portait des sites de fixation de *bantam*. Malheureusement, la répression de ces gènes par RNAi induisait des effets secondaires qui obscurcissaient l'effet du miRNA (en affectant notamment la viabilité des cellules), et le résultat final n'était pas concluant.

Une autre prédiction de cette nouvelle hypothèse pouvait être éprouvée sans nécessiter de nouvelles expériences, simplement en ré-analysant des jeux de données déjà publiés : si effectivement la plupart des cibles prédites sont en réalité des pseudo-cibles, alors leur effet sur le miRNA devrait dépendre de leur niveau d'expression. Un ARNm peu abondant exercerait une faible activité titratrice sur le miRNA, et son interaction avec le miRNA pourrait donc plus facilement se perdre, par mutation, que celle d'un ARNm très abondant. En d'autres termes, les ARNm les plus abondants devraient porter les sites de complémentarité aux miRNA les plus conservés.

D'après la théorie précédente en revanche, l'importance fonctionnelle d'une interaction entre miRNA et ARNm ne devrait pas dépendre du niveau d'expression de l'ARNm (s'il est important de réprimer une cible donnée, elle devrait garder son site de fixation du miRNA, qu'elle soit elle-même faiblement ou fortement exprimée).

J'ai donc recherché d'éventuelles corrélations entre l'abondance des ARNm (dans une variété de tissus de Souris) et la conservation de leurs sites d'interaction avec les miRNA. L'abondance des ARNm a été quantifiée par de nombreuses expériences de transcriptomique : les données étaient donc déjà disponibles. En ce qui concerne la conservation des sites de fixation de miRNA, j'ai choisi de la quantifier par le P_{CT} (*probability of conserved targeting*, définie par Friedman *et al.*, 2009), qui permet de mesurer précisément la pression de sélection qui s'applique sur chaque site de complémentarité aux miRNA de Mammifères.

Le résultat de cette analyse est présenté en figure 1, pour quelques organes ou types cellulaires chez la Souris. Ils montrent que, pour la plupart des miRNA et la plupart des organes et types cellulaires analysés, les cibles prédites de ces miRNA portent des sites de fixation au miRNA d'autant mieux conservés, que ces cibles sont abondantes. Nous avons analysé, en tout, 36 échantillons biologiques, et les 30 *volcano plots* supplémentaires, non montrés en figure 1, ont tous le même aspect, avec un large biais en faveur des corrélations positives.

Cette observation était prédite par l'hypothèse des pseudo-cibles, et n'était pas attendue par la théorie précédente.

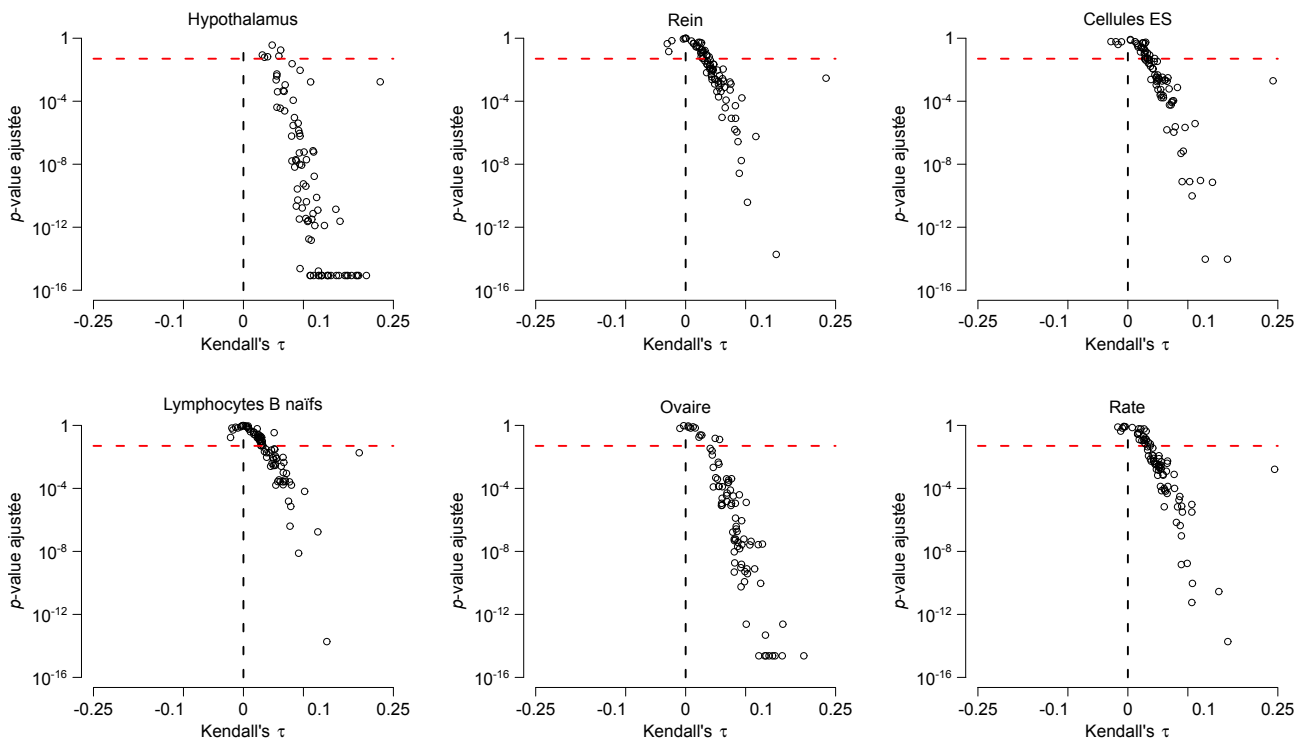


FIGURE 1 – Les ARNm les plus abondants portent les sites d'interaction aux miRNA les plus conservés. Ces graphiques (*volcano plots*) représentent le coefficient de corrélation « Kendall's τ » et sa p -value associée (corrigée pour le test d'hypothèses multiples), qui mesurent la corrélation entre l'abondance des ARNm et la conservation de leurs sites d'interaction aux miRNA. Chaque point représente un miRNA. Les p -values inférieures à $2,2 \times 10^{-16}$ ont été amenées à $2,2 \times 10^{-16}$ pour plus de clarté. La ligne pointillée rouge indique une p -value ajustée de 0,05, et la ligne pointillée noire représente un coefficient de corrélation de 0 (qui indiquerait une absence de corrélation entre abondance des ARNm et conservation de leurs sites de fixation de miRNA).

La mise en œuvre de ce projet de recherche (qui impliquait à la fois des expériences *in vivo*, des analyses informatiques de jeux de données à haut débit, et des mesures de la réponse des réseaux de régulation par des méthodes de biologie des systèmes) nécessitait l'aide de plusieurs personnes, et des financements dédiés. J'ai été averti de l'ouverture d'un appel d'offres pour le recrutement d'une jeune équipe à l'IGH (UPR 1142 du CNRS, à Montpellier), courant 2010. Ma candidature a été retenue (fin 2010), et j'ai pu candidater au financement ATIP-Avenir en décembre 2010, que j'ai obtenu. J'ai donc créé mon équipe à l'IGH en octobre 2011, et initié le contrat ATIP-Avenir en mars 2012.

Entretemps, j'avais recruté une étudiante en thèse (Mlle. Anna Sergeeva, qui venait de terminer son diplôme de master en Russie). Elle a commencé sa thèse sous ma direction au printemps 2011 (donc au LBME, à Toulouse), et m'a suivi à Montpellier à l'automne. Le premier projet d'Anna consistait à rechercher des ARN non-codants, chez la Drosophile, dont le rôle biologique consisterait en la modulation du miRNA *bantam* par titration. Mes analyses informatiques préliminaires avaient en effet identifié 32 loci non-codants, dans le génome de la Drosophile, qui présentaient des sites de complémentarité à *bantam*, sélectivement conservés au cours de l'évolution. Il s'agissait donc de vérifier que ces loci non-codants étaient transcrits en ARN non-codants, pour ensuite invalider ces gènes *in vivo* et mesurer l'effet de ces mutations sur l'activité du miRNA *bantam*.

Parmi ces 32 candidats, Mlle. Sergeeva a testé l'expression de 10 loci, et a montré que 5 de ces 10 loci étaient transcrits en ARN non-codants, clairement détectables chez la puppe de Drosophile.

La caractérisation des extrémités 5' et 3' de ces ARN, par RACE, a posé de nombreux problèmes techniques, et nous n'avons pas pu préciser les limites de ces unités de transcription (nécessaires pour pouvoir ensuite rechercher des mutants de ces loci, ou réprimer ces ARN non-codants par RNAi transgénique). Ce projet reste donc inachevé à l'heure actuelle, il sera repris par une post-doctorante de l'équipe, Natalia Pinzón Restrepo, lorsque son projet actuel sera terminé (*cf* plus bas).

Avec l'aide de Laura Martinez (assistante ingénieure dans l'équipe, recrutée à l'automne 2011), et en collaboration avec le laboratoire de Florence Apparailly (INM, U844 de l'Inserm, à Montpellier), Mlle. Sergeeva a également testé une autre prédiction de l'hypothèse des pseudo-cibles : si effectivement la plupart des cibles prédites de miRNA sont peu sensibles à cette petite répression, alors des individus sauvages devraient présenter une certaine hétérogénéité dans le niveau d'expression de ces gènes (cette hétérogénéité ne leur conférerait pas de faiblesse face à la sélection naturelle, puisque l'activité de ces gènes est robuste vis-à-vis de leur niveau d'expression). En particulier, rien n'interdit que ces fluctuations inter-individus puisse dépasser l'amplitude de la répression guidée par les miRNA.

En revanche, la théorie précédente implique que toutes les cibles prédites de miRNA sont sensibles à leur niveau d'expression (et que c'est leur sensibilité à la répression guidée par le miRNA qui expliquerait la conservation de leur interaction avec le miRNA ; la perte du site de fixation du miRNA abolirait la répression de la cible, et cette petite dérégulation aboutirait à un désavantage sélectif). Une souche sauvage, purifiée par la sélection naturelle, devrait donc avoir éliminé les niveaux d'expression les plus divergents ; notamment, les fluctuations inter-individus dans le niveau d'expression de ces gènes ne devraient pas pouvoir excéder l'amplitude de la répression guidée par les miRNA.

Nous avons donc décidé de mesurer les variabilités inter-individus dans le niveau d'expression des cibles de miR-223 chez la Souris. Ce miRNA est spécifiquement exprimé dans les neutrophiles et leurs progéniteurs, et son effet sur l'expression de ses cibles a été quantifié par des expériences à haut débit (Baek *et al.*, 2008). On connaît ainsi l'effet de miR-223 sur le niveau d'accumulation de ses ARNm-cibles, à l'échelle du transcriptome entier, dans les neutrophiles matures. Il nous restait donc à mesurer la variabilité inter-individus, parmi des souris de souche sauvage, pour l'expression des cibles prédites de miR-223 dans les neutrophiles. Nous avons analysé des souris mâles de la souche C57BL/6 âgées de trois mois, et mesuré l'abondance des ARNm dans leurs neutrophiles par puce à ADN Affymetrix, de manière à utiliser le même matériel biologique et la même procédure expérimentale que Baek *et al.*, 2008.

Nous avons ainsi comparé 5 souris ; nous avons, en outre, mesuré la précision de l'expérience, en analysant simultanément 5 réplicats techniques (voir figure 2A), de façon à soustraire ce bruit technique de la variabilité biologique mesurée. Nos résultats montrent que la variabilité inter-individus est très grande (plus d'un facteur 2 en général), et que, pour la plupart des cibles prédites de miR-223, elle dépassait l'amplitude de la répression guidée par ce miRNA (voir un exemple, sur le gène *Styx*, en figure 2B). Finalement, parmi les 196 cibles prédites de miR-223, l'amplitude de la répression par miR-223, et la variabilité inter-individus, avaient été mesurées sans ambiguïté pour 189 cibles. Parmi ces 189 gènes, 168 présentent une variabilité inter-individus qui dépasse leur répression guidée par miR-223 (avec $p < 0,05$). En d'autres termes, pour environ 90 % des cibles prédites de miR-223, nous montrons que ces gènes semblent trop robustes pour être fonctionnellement affectés par miR-223 dans les neutrophiles : leur expression fluctue davantage entre souris sauvages, sans leur conférer de phénotype néfaste.

À l'automne 2012, Mlle. Sergeeva a décidé d'arrêter sa thèse, pour rentrer en Russie suite à des problèmes personnels. Natalia Pinzón Restrepo (post-doctorante arrivée dans l'équipe le 1^{er} mars 2012) a, quant à elle, pris en charge le test d'une autre prédiction de l'hypothèse des pseudo-cibles : la comparaison de l'abondance de miRNA et de leurs cibles prédites, dans une population cellulaire homogène. L'hypothèse des pseudo-cibles implique que, dans les cellules où la titration du miRNA est bénéfique, l'abondance cumulée des cibles (pondérée par leur nombre de sites de fixation du miRNA) est, au moins, de l'ordre de grandeur de l'abondance du miRNA. La théorie précédente,

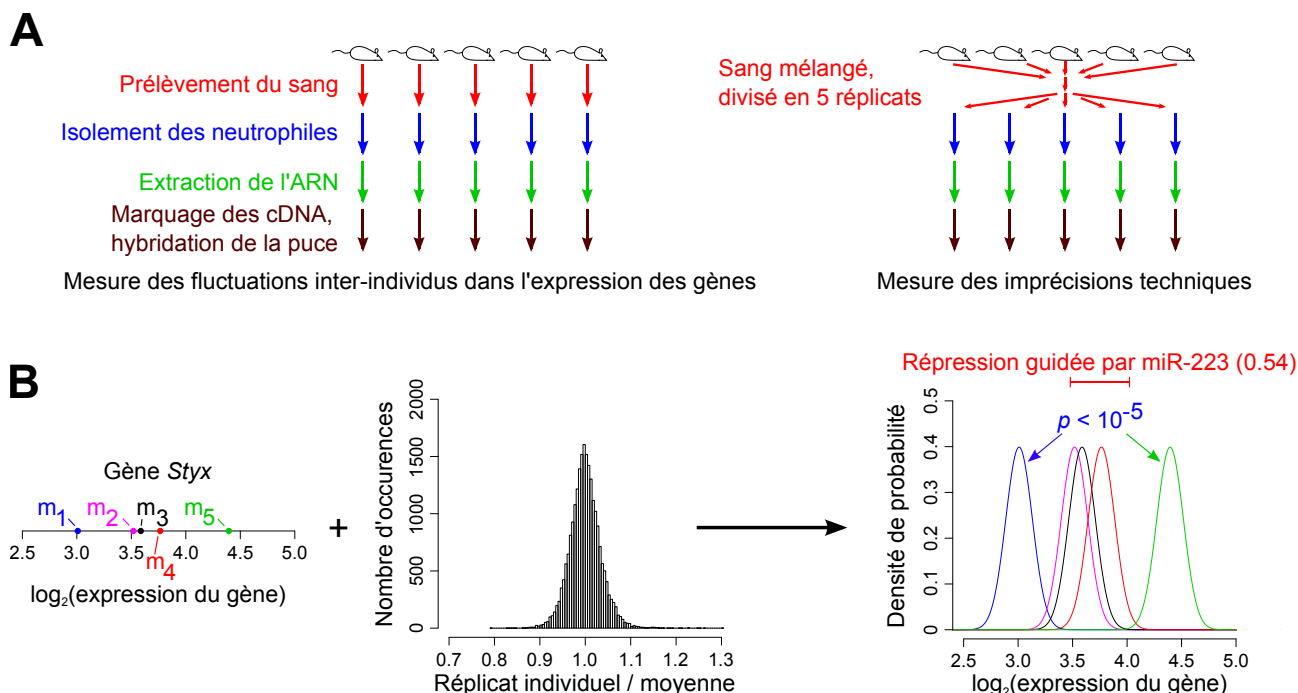


FIGURE 2 – Mesure de la variabilité inter-individus dans l’expression des cibles de miR-223. **A.** Représentation schématique de la procédure expérimentale. **B.** Résultat représentatif (sur l’exemple du gène *Styx*). Les mesures d’expression du gène sur les 5 souris (m_1 à m_5 , panneau de gauche) et la mesure des imprécisions techniques, mesurées sur les sondes présentant une intensité de signal similaire, parmi les 5 réplicats techniques (panneau du milieu) nous ont permis de calculer la densité de probabilité de la distribution de la valeur biologique réelle, pour chaque souris (panneau de droite). Nous avons utilisé cette distribution pour calculer la probabilité que la variabilité entre les 5 souris soit plus petite que l’amplitude de la répression guidée par miR-223 sur ce gène (qui est représentée par une barre horizontale rouge sur le panneau de droite; d’après Baek *et al.*, 2008, $\log_2(\text{répression})=0,54$ pour *Styx*). Pour le gène *Styx*, la variabilité inter-individus est donc supérieure à la répression guidée par miR-223 dans les neutrophiles, et il est très peu probable que cette différence soit due aux imprécisions techniques dans la mesure de l’expression du gène ($p < 10^{-5}$ pour *Styx*).

qui estime que la seule fonction des interactions entre ARNm et miRNA consiste en la répression de l’ARNm, implique que le miRNA est en excès par rapport à ses cibles (Garcia *et al.*, 2011 ; Wee *et al.*, 2012). Nous avons donc décidé de quantifier, en termes absolus (le résultat devait être un nombre de molécules par cellule), l’abondance des miRNA miR-1, miR-133 et miR-206 au cours de la différenciation des cellules C2C12. Il est possible d’induire la différenciation de cette lignée cellulaire murine (dérivée de myoblastes) en myotubes en changeant les conditions de culture, et l’expression de ces trois miRNA spécifiques des myotubes est progressivement activée au cours de la différenciation (Chen *et al.*, 2006 ; Kim *et al.*, 2006).

Mlle. Pinzón Restrepo a donc induit la différenciation de cellules C2C12, en trois réplicats, et mesuré l’abondance de miR-133 d’une part, et de miR-1 et de miR-206 d’autre part, par Northern blot (les séquences de miR-1 et de miR-206 sont trop similaires pour permettre une distinction fiable par Northern blot ; comme la liste de cibles prédites pour ces deux miRNA est identique, nous avons donc décidé de ne pas les distinguer). En calibrant les membranes par des oligonucléotides ARN synthétiques, de mêmes séquences que les miRNA à quantifier, nous avons pu estimer précisément l’abondance de ces miRNA au cours de la différenciation des cellules C2C12 (voir figure 3A). Nous avons mesuré le nombre de cellules analysées dans chaque échantillon, par cytométrie de flux, ce qui nous a permis de calculer le nombre de molécules de miR-133, et de molécules de miR-1 + miR-206, par cellule (voir figure 3B).

Nous sommes actuellement en train de mesurer l'abondance des ARNm sur les trois réplicats du jour 0, du jour 3, et du jour 6 de différenciation. Des banques de séquençage par RNA-Seq pour ces 9 échantillons d'ARN sont en cours de préparation par le Beijing Genomic Institute (Chine). Afin de convertir ces données de RNA-Seq en mesures absolues, nous avons introduit dans chaque échantillon une collection de transcrits *in vitro*, en quantité contrôlée. Ces transcrits (au nombre de 27) n'ont pas d'homologie avec le génome de la Souris (pour ne pas être confondus avec les ARN des cellules C2C12), ils ont été purifiés sur gel d'acrylamide (pour éliminer les transcrits incomplets), ils mesurent entre 400 et 800 nt, et ils sont polyadénylés en 3'. Le protocole classique de RNA-Seq commence par une étape de purification des ARN polyadénylés, mais il est possible de s'affranchir de cette étape (en la remplaçant par une étape d'élimination sélective des ARN ribosomiques). Nous avons comparé ces deux méthodes sur un échantillon-pilote (voir figure 3C et D), ce qui nous a également permis de nous assurer que les transcrits *in vitro* étaient introduits en quantité raisonnable (suffisante pour qu'ils soient précisément quantifiés, mais pas excessive : ils n'occupent que 8 % de la profondeur de séquençage). Le protocole poly(A)-dépendant fournit une plus grande profondeur de séquençage, et, partant, une meilleure précision sur la régression reliant l'abondance réelle des ARN, au nombre de leurs reads détectés par RNA-Seq. Nous avons donc choisi de séquencer les 9 échantillons par ce protocole.

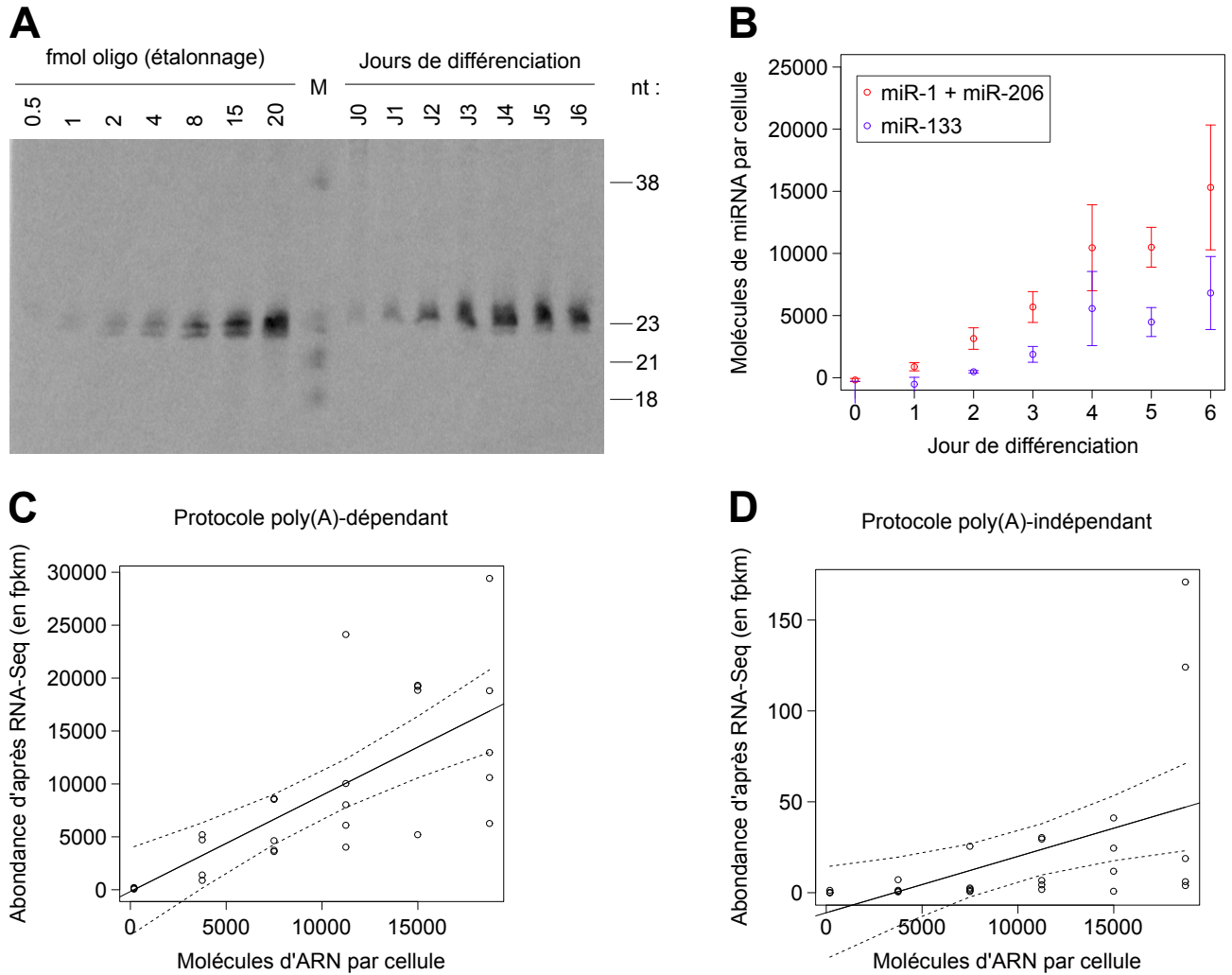


FIGURE 3 – Quantification de miR-1, miR-206 et miR-133 et des ARNm au cours de la différenciation des cellules C2C12. **A.** Quantification de miR-1 et miR-206 par Northern blot au cours de la différenciation. Un mélange équimolaire d'oligonucléotides synthétiques portant les séquences de miR-1 et de miR-206 a été déposé en quantité croissante sur le gel, pour étalonner le Northern blot (7 pistes de gauche). M : marqueur de taille (18 ; 21 ; 23 et 38 nt). 7 pistes de droite : 20 μ g d'ARN total de cellules C2C12 après 0, 1, 2, 3, 4, 5 et 6 jours de différenciation, respectivement. **B.** Nombre de molécules de (miR-1 + miR-206), et de miR-133, par cellule au cours de la différenciation (valeurs issues de trois réplicats de la différenciation, analysés chacun comme sur le panneau **A**). Les barres d'erreur indiquent l'erreur-type parmi les trois réplicats. **C et D.** Résultat de la quantification des ARNm d'un échantillon-pilote (l'un des trois réplicats du 6ème jour de différenciation), par le protocole classique de RNA-Seq (panneau **C**) et par un protocole qui n'implique pas de purification des ARN polyadénylés (panneau **D**). Chaque point représente l'un des 27 transcrits *in vitro* introduits dans l'échantillon. (de 187 à 18700 molécules par cellule, selon les transcrits). L'abondance de chaque ARN est exprimée en *fpkm* (*fragments per kb and per million reads*), unité qui normalise l'abondance des *reads* à la fois pour la longueur des ARN (parce que les ARN les plus longs génèrent davantage de *reads*, à quantité molaire égale) et pour la profondeur de séquençage de la banque. La ligne en trait plein représente la régression linéaire du nombre de *fpkm* en fonction de la quantité de transcrit introduite, pour les 27 transcrits *in vitro*; les lignes en pointillés représentent les limites de l'intervalle de confiance à 95 % où se situeraient des points expérimentaux supplémentaires.

2.2 Histoire évolutive des petits ARN régulateurs chez les Métazoaires

À la fin de mon stage post-doctoral dans le laboratoire du professeur P. Zamore (University of Massachusetts Medical School, États-Unis), j'avais initié un projet visant à décrire le répertoire de petits ARN régulateurs chez l'Anémone de mer *Nematostella vectensis*. C'est un Cnidaire, dont le dernier ancêtre commun avec les Bilatériens¹ a plus de 600 millions d'années. Nous avons donc séquencé les petits ARN de neuf stades développementaux de *Nematostella* (œuf non fécondé, blastula, gastrula, planula précoce, planula tardive, métamorphose, polype primaire, mâle adulte, femelle adulte), en prenant soin de diviser chaque échantillon en deux : la première moitié a servi à préparer une banque de Small RNA-Seq selon le protocole classique, l'autre moitié a d'abord été oxydée par le periodate de sodium, en présence d'acide borique, de manière à éliminer les petits ARN qui ne portent pas de modification chimique sur leur extrémité 3' (ce protocole enrichit donc les banques en petits ARN 2'-O-méthylés sur leur 3', tels que les piRNA chez les Bilatériens).

Depuis mon arrivée au CNRS en janvier 2009, j'ai analysé ces banques de séquençage, pour annoter les miRNA de *Nematostella*. Dans l'intervalle, un autre laboratoire avait réalisé une expérience similaire (mais en ne séquençant qu'un échantillon d'ARN de *Nematostella*, extrait d'un mélange de stades développementaux ; Grimson *et al.*, 2008). Notre expérience, plus détaillée, nous a permis de compléter la liste de miRNA décrite par Grimson *et al.*, 2008, et de revisiter quelques-uns des petits ARN qu'ils avaient annoté. La figure 4 illustre nos résultats (le panneau **A** montre un miRNA précédemment décrit, et que nous avons confirmé ; le panneau **B**, un nouveau miRNA ; et le panneau **C**, un ARN précédemment décrit comme un miRNA, mais qui s'avère être un siRNA).

J'ai, d'autre part, lancé une collaboration avec le laboratoire du professeur U. Technau, à l'université de Vienne (Autriche) : ce laboratoire est spécialisé dans l'étude du développement de *Nematostella*, et il a pu établir la localisation de plusieurs de ces miRNA, et de leurs cibles prédites, par des expériences d'hybridation *in situ* de sondes fluorescentes. En partageant les analyses informatiques et les expériences entre le laboratoire d'U. Technau, le laboratoire de P. Zamore (mon ancien laboratoire de stage post-doctoral), et moi-même, nous avons réalisé une analyse du « dégradome » chez les polypes primaires de *Nematostella*. Cette expérience est une variante du RNA-Seq, qui permet de ne séquencer que les fragments d'ARN qui portent un monophosphate en 5'. Elle nous a permis de montrer que les miRNA de *Nematostella* guident fréquemment le clivage endonucléolytique d'ARN-cibles, à la manière des siRNA, et de plusieurs miRNA chez les Plantes (cette réaction de clivage libère un fragment 3' dont l'extrémité 5' porte un monophosphate : le séquençage du dégradome permet donc de détecter les produits 3' de cette réaction de clivage).

À la différence des miRNA de Bilatériens, les miRNA de ce Cnidaire guident donc fréquemment le clivage de leurs cibles : la similarité de ce mode de régulation entre les Plantes et les Cnidaires suggère qu'il s'agit du mode de régulation ancestral par les miRNA, et que la répression traductionnelle et la dégradation exonucléolytique (mécanismes les plus courants de la répression des cibles de miRNA chez les Bilatériens) pourrait donc être une innovation récente chez les Animaux. Tous ces travaux sont présentés dans un manuscrit, actuellement en révision à *Genome Research*.

Enfin, ces résultats ont également mis en évidence des siRNA endogènes chez *Nematostella*. Outre les piRNA (déjà décrits par Grimson *et al.*, 2008, confirmés par nos analyses) et les miRNA, *Nematostella* possède donc la troisième classe de petits ARN régulateurs connus chez les Bilatériens. Bien que non discutées dans le manuscrit soumis, ces conclusions présentent un intérêt supplémentaire : chez *Nematostella*, au même titre que chez les Bilatériens, les siRNA et les piRNA ont tendance à être complémentaires des séquences de transposons. Le rôle de ces deux familles de petits ARN dans la répression des éléments transposables a été établi chez les Bilatériens (voir par exemple Brennecke *et al.*, 2007 ; Aravin *et al.*, 2007 ; Ghildiyal *et al.*, 2008 ; Kawamura *et al.*, 2008 ; Chung *et al.*, 2008 ; Tam *et al.*, 2008 ; Watanabe *et al.*, 2008). La redondance fonctionnelle de ces deux classes de petits ARN semble donc très ancienne, et, de manière surprenante, elle a été préservée chez de nombreuses

1. Auxquels appartiennent la plupart des organismes-modèles animaux habituels : Insectes, Nématodes, Mollusques, Vertébrés.

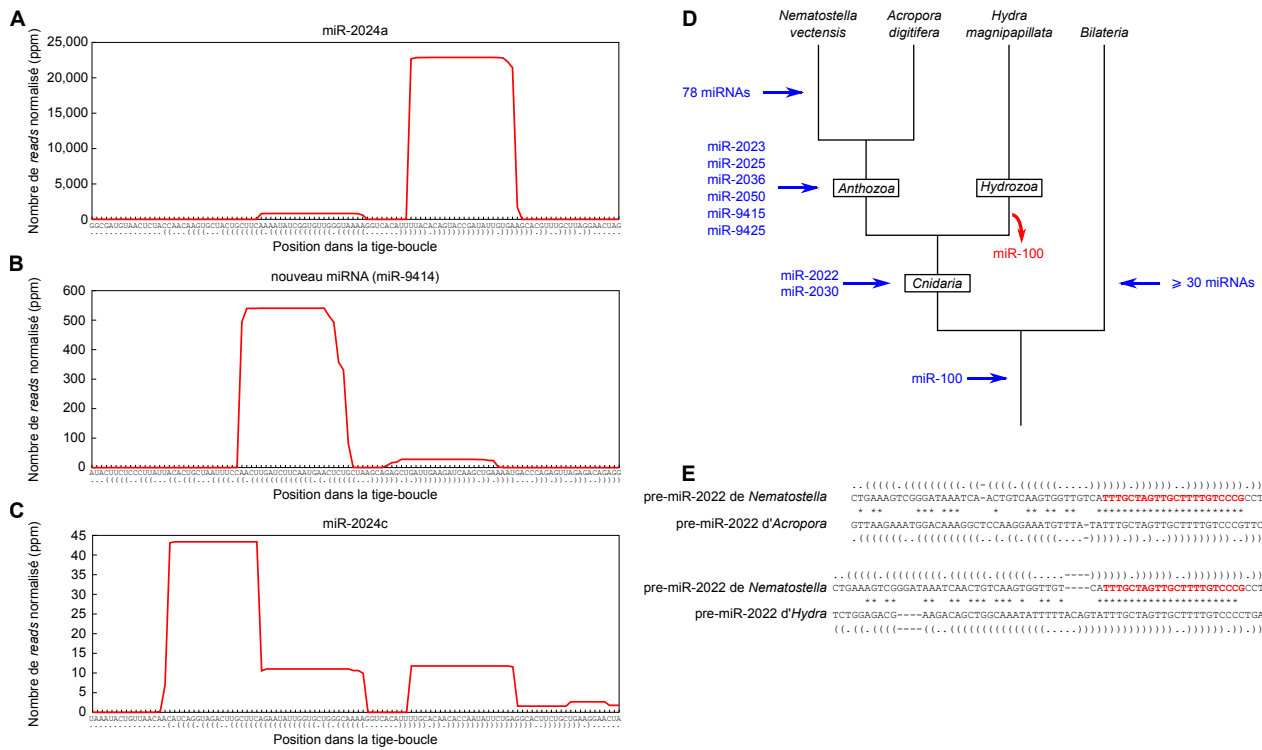


FIGURE 4 – Les miRNA de *Nematostella vectensis*. **A, B et C.** Profil d’abondance des *reads* détectés dans l’ensemble des 18 banques de Small RNA-Seq, le long de la séquence du précurseur en tige-boucle d’un miRNA précédemment décrit par Grimson *et al.*, 2008 (miR-2024a, panneau **A**), d’un miRNA qui n’était pas connu (miR-9414, panneau **B**) et d’un petit ARN que Grimson *et al.*, 2008 avaient décrit comme un miRNA, mais qui s’avère être un siRNA, issu d’une tige-boucle débitée en une variété de petits ARN de séquences diverses (miR-2024c, panneau **C**). **D.** Scénario évolutif probable pour les miRNA du lignage de *Nematostella*. **E.** Alignement des séquences et des structures secondaires prédites pour pre-miR-2022 chez quelques Cnidaires modèles. La séquence de miR-2022 mature est indiquée en rouge. Les nucléotides conservés sont annotés avec une astérisque, les tirets indiquent des délétions dans l’alignement. La structure secondaire est représentée par des points (pour les nucléotides non-appariés) et des parenthèses (pour les nucléotides appariés).

espèces actuelles. Le maintien de cette redondance apparente dans une telle variété d’espèces, pendant plus d’un demi-milliard d’années, suggère que chacune des deux voies est nécessaire, même en présence de l’autre. Soit une contrainte encore inconnue empêche les siRNA de réprimer les transposons contrôlés par les piRNA, et les piRNA de réprimer les transposons contrôlés par les siRNA ; soit chacune de ces deux classes joue une autre fonction, que ne peut pas remplir l’autre classe (par exemple, la répression de gènes non-transposables ; mais il semble que piRNA et siRNA en sont tous les deux capables : Ghildiyal *et al.*, 2008 ; Okamura *et al.*, 2008 ; Rajasethupathy *et al.*, 2012).

2.3 Déterminants structuraux de la biogenèse des microARN

Mon collègue le professeur Y. Tomari est un biochimiste de l'université de Tōkyō, qui travaille sur la biogenèse et le mode d'action des petits ARN régulateurs. Il a sollicité mon aide pour des analyses bio-informatiques qui allaient compléter les résultats des expériences de biochimie réalisées dans son laboratoire.

Les miRNA sont produits sous la forme d'un petit duplex, où ils sont appariés à un autre petit ARN, de séquence imparfaitement complémentaire, appelé « miRNA* ». C'est ce duplex qui est chargé sur une protéine de la famille Argonaute (les protéines effectrices de la répression des cibles, guidées par les petits ARN). Après chargement, le duplex miRNA/miRNA* est ouvert, et le brin miRNA* est éliminé. Les expériences du laboratoire du professeur Tomari ont montré qu'un mésappariement situé dans le tiers 5' du miRNA, ou dans le début de la deuxième moitié de sa séquence, facilitaient cette étape de la maturation du complexe effecteur. J'ai donc analysé, sur l'ensemble des miRNA connus chez la Drosophile, le Nématode et la Souris, la structure des duplex miRNA/miRNA*, et montré que la grande majorité d'entre eux avaient une structure qui optimisait l'ouverture du duplex (Kawamata *et al.*, 2009).

D'autre part, l'enzyme responsable de la production de ce duplex à partir de son dernier précurseur (l'enzyme Dicer-1 chez la Drosophile) est capable de reconnaître spécifiquement les précurseurs de miRNA (des tiges-boucles d'environ 70 nt chez les Animaux, appelées « pre-miRNA »). Le laboratoire du professeur Tomari a recherché l'origine de cette spécificité, et a montré que le domaine hélicase de Dicer-1 en était responsable, en interagissant physiquement avec plusieurs éléments structuraux du pre-miRNA. J'ai vérifié la généralité de ces règles, en analysant systématiquement la taille des différents éléments des pre-miRNA naturels de Drosophile (Tsutsumi *et al.*, 2011).

2.4 Le rôle du nucléotide 5' des microARN

Au cours de mon stage post-doctoral, je m'étais aperçu (par des expériences de biochimie) que l'identité du premier nucléotide d'un miRNA affectait l'efficacité de chargement du duplex miRNA/miRNA* sur les protéines Argonaute chez la Drosophile. Ces travaux étaient incomplets : je souhaitais vérifier leur généralité, par des analyses informatiques, avant de les publier.

C'est en France, alors que je travaillais au LBME à Toulouse (2009-2011), que j'ai terminé ce projet. J'ai pu montrer que cette préférence sur le nucléotide en 5' renforçait l'asymétrie fonctionnelle entre le miRNA et le miRNA* : c'est lors du chargement du duplex sur la protéine Argonaute qu'est déterminée l'identité du brin qui restera stablement associé à la protéine (le brin miRNA), et celle du brin qui sera dégradé (le brin miRNA*). Une uridine en 5' d'un brin d'ARN facilite le chargement du duplex dans l'orientation qui fera de ce brin le miRNA ; or la plupart des miRNA de Drosophile, de Nématode et de Souris portent une uridine en 5', à la différence de leurs miRNA*. Ce biais de séquence contribue donc à garantir que le duplex miRNA/miRNA* donnera un complexe effecteur très homogène, dont le petit ARN-guide sera le plus souvent le brin miRNA. Des analyses plus détaillées ont de plus montré que l'ordre de préférence sur le premier nucléotide dépendait de la nature du reste de la séquence, et notamment, de l'identité du deuxième nucléotide.

Ces résultats montrent donc que le dinucléotide en 5' des miRNA et miRNA* contribue à l'asymétrie du chargement du duplex, et donc, à l'homogénéité du complexe effecteur (Seitz *et al.*, 2011).

2.5 Collaborations avec des partenaires académiques

J'ai participé à deux collaborations non contractuelles, avec des laboratoires étrangers (les laboratoires des professeurs Technau, en Autriche, et Zamore, aux États-Unis : voir section 2.2 ; et du professeur Tomari, au Japon : voir section 2.3). Ces collaborations ont respectivement abouti à un manuscrit en cours de révision, et à deux publications (voir détail plus haut).

Je suis également impliqué, en tant que partenaire, dans un projet coordonné par Denis Tagu (INRA de Rennes) et financé par l'ANR (projet « miRNAadapt »). Ce projet vise à explorer le rôle des miRNA dans la plasticité phénotypique du Puceron, dont le mode de vie et de reproduction varient en fonction de stimuli saisonniers. Nos premiers résultats suggèrent que certaines cibles de miRNA sont régulées différemment au cours de l'année, et nous sommes en train de chercher les relations régulatrices responsables de ce phénomène.

2.6 Place de ma recherche dans celle de mon unité, mobilité géographique

Entre janvier 2009 et septembre 2011, j'étais membre de l'équipe de Jérôme Cavaillé, au LBME (UMR 5099, CNRS et université Paul Sabatier, à Toulouse). Cette unité (<https://www-lbme.biotoul.fr>) a une longue tradition d'expertise dans le domaine des petits ARN non-codants. Outre mes travaux propres (*cf* plus haut), j'ai contribué à l'activité de l'équipe Cavaillé par des analyses bio-informatiques (l'équipe avait généré des souris mutantes pour un locus de gènes de miRNA et en analysait le phénotype, notamment par des expériences de microarray ; j'ai identifié les gènes différemment exprimés après cette délétion, et recherché les cibles prédites des miRNA en question). Cette contribution informelle (les résultats ne seront pas nécessairement inclus dans le manuscrit préparé par le laboratoire) pourrait augmenter si les besoins de la publication l'exigent.

J'ai changé d'unité de rattachement en octobre 2011, pour rejoindre l'IGH (UPR 1142 du CNRS, à Montpellier), après avoir été retenu pour créer une nouvelle équipe dans cet institut.

L'IGH (<http://www.igh.cnrs.fr/FR/index.php>) est structuré en trois départements (« Dynamique du génome », « Bases moléculaires de pathologies humaines » et « Génétique et développement »). Mon équipe appartient au département « Génétique et développement », et les questions que nous traitons, sur le rôle des miRNA à l'échelle de l'organisme, nous amènent à interagir fréquemment avec nos collègues développementalistes. D'autre part, plusieurs équipes de l'institut travaillent sur les petits ARN régulateurs (les piRNA, par les équipes de Martine Simonelig et de Séverine Chambeyron ; les miRNA, par l'équipe de Moncef Benkirane ; les siRNA, par l'équipe de Rosemary Kiernan).

Enfin, d'un point de vue méthodologique, nos travaux nous rapprochent des nombreuses équipes qui s'intéressent aux analyses à haut débit. Nous organisons des réunions mensuelles d'échange méthodologique, entre bio-informaticiens de l'institut (outre mon équipe, celles de Giacomo Cavalli, de Marie-Paule Lefranc, de Philippe Pasero, de Séverine Chambeyron, de Marcel Méchali, et de Martine Simonelig, participent activement). Notre groupe de bio-informaticiens de l'IGH a récemment commencé à dispenser des formations à la bio-informatique et aux statistiques, pour nos collègues biologistes (voir <http://www.igh.cnrs.fr/equip/Seitz/equipe-tutoriels.html> pour ma contribution).

L'expertise de mon équipe dans les analyses bio-informatiques nous a également amené à collaborer avec une autre équipe de l'institut, celle de Marcel Méchali, sur deux projets de cette équipe qui nécessitaient des analyses à haut débit. L'un de ces projets est actuellement soumis pour publication, l'autre est en cours de rédaction.

2.7 Distinction scientifique

La prime d'excellence scientifique m'a été attribuée en décembre 2010.

2.8 Publications scientifiques des 10 derniers semestres

2.8.1 Revues à comité de lecture

1. Seitz, H. (2009). Redefining microRNA targets. *Curr Biol*, 19(10) : 870–873. PDF de l'article
2. Kawamata, T., Seitz, H., et Tomari, Y. (2009). Structural requirement of miRNAs for RISC-loading and Slicer-independent unwinding. *Nat Struct Mol Biol*, 16(9) : 953–960. PDF de l'article
3. Tsutsumi, A., Kawamata, T., Izumi, N., Seitz, H., et Tomari, Y. (2011). Recognition of the pre-miRNA structure by *Drosophila* Dicer-1. *Nat Struct Mol Biol*, 18(10) : 1153–1158. PDF de l'article
4. Klattenhoff, C., Xi, H., Li, C., Lee, S., Xu, J., Khurana, J. S., Zhang, F., Schultz, N., Koppetsch, B. S., Nowosielska, A., Seitz, H., Zamore, P. D., Weng, Z., et Theurkauf, W. E. (2009). The *Drosophila* HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters. *Cell*, 138(6) : 1137–1149. PDF de l'article
5. Li, C., Vagin, V. V., Lee, S., Xu, J., Ma, S., Xi, H., Seitz, H., Horwich, M. D., Syrzycka, M., Honda, B. M., Kittler, E. L., Zapp, M. L., Klattenhoff, C., Schulz, N., Theurkauf, W. E., Weng, Z., et Zamore, P. D. (2009). Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies. *Cell*, 137(3) : 509–521. PDF de l'article
6. Sergeeva, A. M., Pinzón Restrepo, N., et Seitz, H. (2013). Quantitative aspects of RNA silencing in metazoans. *Biochemistry (Moscow)*, 78(6) : 613–626. PDF de l'article
7. Seitz, H., Tushir, J. S., et Zamore, P. D. (2011). A 5'-uridine amplifies miRNA/miRNA* asymmetry in *Drosophila* by promoting RNA-induced silencing complex formation. *Silence*, 2 : 4. PDF de l'article
8. Ghildiyal, M., Xu, J., Seitz, H., Weng, Z., et Zamore, P. D. (2010). Sorting of *Drosophila* small silencing RNAs partitions microRNA* strands into the RNA interference pathway. *RNA*, 16(1) : 43–56. PDF de l'article
9. Seitz, H. (2010). siRNAs : the hidden face of the small RNA world. *Curr Biol*, 20(3) : R108–R110. PDF de l'article

2.8.2 Conférences invitées dans des congrès

1. Conférence de 30 minutes (titre : « Rethinking the biological function of small RNAs in animals ») au *Microsymposium on small RNAs* à Vienne (Autriche) le 19 mai 2009.
2. Conférence d'1 heure (titre : « Butterfly effect in molecular biology : the case of microRNAs ») à la *Schlumberger Foundation conference "Small Silencing RNA Biology and Mechanism"* à Tourtour (France) 24 avril 2012.
3. Conférence de 30 minutes (titre : « miRNA-guided repression vs. biological robustness : revisiting the significance of miRNA-mediated regulation ») au *BTCure miRNA workshop* à Montpellier (France) le 13 octobre 2012.
4. Conférence d'1 heure (titre : « Systemic impact of microRNAs in animals ») à l'*International course on Non-coding genome* de l'Institut Curie le 11 décembre 2012.
5. Conférence de 45 minutes (titre : « Statistical analysis of small RNA-seq data ») au symposium *Statistical analysis of RNA-seq data : advances and challenges* de l'Institut Pasteur le 26 novembre 2013.
6. Conférence d'1 heure 30 (titre : « Annotation des ARN non codants chez les Eucaryotes — les enjeux biologiques ») à la journée *DE^CODAGE* du GdR BiM le 26 avril 2013.
7. Conférence d'1 heure (titre : « Petits ARN régulateurs : la génétique des années 2000 ») au congrès annuel de la Nouvelle Société française d'athérosclérose le 6 juin 2009.

8. Conférence de 30 minutes (titre : « Biological function of microRNAs in *Drosophila* ») à l'*Annual French conference on Drosophila* à Clermont-Ferrand (France) le 12 septembre 2012.
9. Conférence de 30 minutes (titre : « The biology of microRNAs in animals ») au *Fifth European workshop on immune-mediated inflammatory diseases* à Sitges (Espagne) le 3 décembre 2010.
10. Conférence de 30 minutes (titre : « The genomics of small RNAs in insects ») au *Fifth annual Arthropod genomics symposium* à Kansas City (États-Unis) le 10 juin 2011.

2.8.3 Actes de colloques à comité de lecture

(aucun)

2.8.4 Publications dans des revues sans comité

1. Seitz, H. (2010) Les statistiques en biologie expérimentale : pourquoi ? Comment ? (Première partie) *Regard sur la biochimie*, juillet 2010 : 7-9. PDF du numéro de juillet 2010 de *Regard sur la biochimie*
2. Seitz, H. (2010) Les statistiques en biologie expérimentale : pourquoi ? Comment ? (Seconde partie) *Regard sur la biochimie*, octobre 2010 : 7-9. PDF du numéro d'octobre 2010 de *Regard sur la biochimie*
3. Pinzón Restrepo, N., Martinez, L., Seitz, H. et Cavaillé, J. (2013) Les petits ARN entrent dans l'arène. *Pour la science*, dossier n°81 « L'hérédité sans gènes » (octobre-décembre 2013). Sommaire du numéro
4. Seitz, H. (2010) Qu'est-ce qu'une cible de microARN ? *Regard sur la biochimie*, juillet 2010 : 5-7. PDF du numéro de juillet 2010 de *Regard sur la biochimie*

2.8.5 Communications à des congrès, symposiums

1. Conférence de 10 minutes (titre : « Every predicted miRNA target is not functionally targeted : revisiting the nature of miRNA/mRNA interaction ») au *Keystone Symposium "RNA silencing (C7)"* à Whistler (Canada) le 22 mars 2013.
2. Conférence de 12 minutes (titre : « A new perspective on miRNA-mediated regulation ») au *Sixteenth annual meeting of the RNA society* à Kyōto (Japon) le 15 juin 2011.
3. Conférence de 10 minutes (titre : « The spectacular diversity of small silencing RNAs in animals is ancient and not a vertebrate innovation ») au *Keystone Symposium "The biology of RNA silencing (D6)"* à Victoria (Canada) le 28 avril 2009.
4. Conférence de 15 minutes (titre : « Systemic impact of microRNA-mediated regulation in animals ») au cinquième congrès du Tōkyō RNA club, Tōkyō (Japon) le 13 juin 2011.

2.8.6 Séminaires, *workshops*

1. Année 2013 : séminaires donnés à Montpellier (12 février), Oxford (3 mai), Heidelberg (24 juillet) et Toulouse (18 octobre).
2. Année 2012 : séminaires donnés à Rennes (16 mars), Grenoble (13 juin), Montpellier (29 juin) et Roscoff (3 décembre).
3. Année 2011 : séminaires donnés à Copenhague (2 mai), Toulouse (30 mai) et Montpellier (17 novembre).
4. Année 2010 : séminaires donnés à Montpellier (9 avril), Évry (6 mai), Paris (7 mai), Munich (15 octobre) et Strasbourg (10 décembre).

2.8.7 Livres et ouvrages

(aucun)

2.8.8 Chapitres d'ouvrages

1. Renalier, M.-H., Tirmarche, S., et Seitz, H. (2011). *MicroRNA : Expression, Detection and Therapeutic Strategies*, chapter “Biological functions of microRNAs in animals”. Nova Science Publishers. ISBN 978-1-61122-846-5 PDF du chapitre

2.8.9 Logiciels

(aucun)

2.8.10 Autres

(aucun)

3 Enseignement, formation et diffusion de la culture scientifique

3.1 Thèse et post-doctorats

J'ai encadré, à partir de mars 2011, la thèse de Mlle. Anna Sergeeva. Le sujet de Mlle. Sergeeva portait sur la mesure de l'effet de titration des miRNA *in vivo*, et de la robustesse de l'activité biologique des cibles prédites de miRNA vis-à-vis des variations de leur niveau d'expression. Elle a pu mettre en évidence l'existence de 5 ARN non-codants, portant des sites de complémentarité phylogénétiquement conservés au miRNA *bantam*, et montrer que, dans les neutrophiles de Souris, 90% des cibles prédites pour le miRNA miR-223 semblaient trop robustes pour être fonctionnellement affectées par le miRNA (voir section 2.1 de la partie « Recherche scientifique »). Suite à des problèmes personnels, Mlle. Sergeeva a décidé d'interrompre sa thèse et de rentrer dans son pays d'origine, la Russie, en novembre 2012. Elle travaille désormais pour le bureau de la compagnie *New England Biolabs* à Moscou.

J'encadre le post-doctorat de Mlle. Natalia Pinzón Restrepo depuis mars 2012. Mlle. Pinzón Restrepo a mesuré l'abondance de miRNA et des ARNm (en mesure absolue : elle a déterminé le nombre de leurs molécules par cellule) au cours de la différenciation des cellules C2C12 (une lignée cellulaire murine, dont on peut induire la différenciation en myotubes en culture). Elle est également impliquée dans deux collaborations de notre équipe (avec le laboratoire de M. Méchali, à Montpellier, et avec le laboratoire de M. Cock, à Roscoff), où ses travaux consistent respectivement à analyser des données génomiques issues d'expériences à haut débit chez le Nématode, et à confirmer expérimentalement l'existence de petits ARN chez l'Algue brune *Ectocarpus siliculosus*. Mlle. Pinzón Restrepo est financée par une allocation de la Ligue contre le cancer depuis mars 2013, pour une durée de deux ans renouvelable un an. La qualité du travail de Mlle. Pinzón Restrepo lui a valu de présenter ses résultats oralement au *Cold Spring Harbor Laboratory Meeting "Regulatory and non-coding RNAs"* en août et septembre 2012, puis au congrès SifrARN de la Société française de biochimie et de biologie moléculaire, en novembre 2013.

Depuis novembre 2013, je co-encadre le post-doctorat de M. Blaise Li (co-encadré avec ma collègue Séverine Chambeyron, qui dirige une autre équipe de l'IGH). M. Li analyse la conservation des sites de complémentarité aux miRNA chez les Animaux, pour généraliser les résultats montrés en figure 1, page 8.

3.2 Participation à l'enseignement

J'ai dispensé des cours sur les petits ARN régulateurs eucaryotiques (miRNA, siRNA et piRNA) dans divers établissements au cours des 10 derniers semestres :

- Cours de trois heures pour le M2R « Gènes, cellules et développement » de Toulouse (16 octobre 2009).
- Cours de deux heures pour les étudiants de L3 du cursus « chimie-biologie » de l'École normale supérieure (7 mai 2010).
- Cours de trois heures pour l'école doctorale de biologie de Strasbourg (10 janvier 2011).
- Cours de deux heures aux étudiants de master 1 de Rennes (20 novembre 2012).
- Cours de deux heures aux étudiants de master 2 au Muséum national d'histoire naturelle à Paris (21 novembre 2012).
- Cours de deux heures aux étudiants de master 1 à Lyon (10 décembre 2012).
- Cours de deux heures pour les étudiants de L3 de la formation interuniversitaire de biologie de l'École normale supérieure (28 février 2013).

Outre ces cours généraux sur les petits ARN, j'ai été amené à faire des cours plus spécifiques sur mes projets de recherche, ou sur des méthodologies que j'utilise fréquemment :

- Cours de présentation des activités du laboratoire aux étudiants de master 1 de Toulouse (durée : 1 heure ; 10 décembre 2009).
- Sessions d’initiation à l’utilisation d’Unix et à la bio-informatique à destination du personnel de l’IFR 109 (Toulouse ; printemps et été 2009) et de l’IGH (Montpellier ; décembre 2013). Durée : 5 à 6 heures par établissement.
- Cours « Les statistiques en biologie moléculaire », à destination du personnel de l’IFR 109 (Toulouse ; printemps et été 2010), également donné à l’University of Massachusetts Medical School en mars 2010, et à l’IGH (Montpellier) en avril 2010 (durée de chaque cours : 1 heure). Une version abrégée de ce cours est disponible en deux articles dans les numéros de juillet et octobre 2010 de *Regards sur la Biochimie* : pages 7 à 9 de du numéro de juillet 2010, et pages 7 à 9 du numéro d’octobre 2010 : <http://sfbbm.fr/index.php/publications/regard-sur-la-biochimie2>
- Conférence d’1 heure (titre : « Systemic impact of microRNAs in animals ») à l’*International course on Non-coding genome* de l’Institut Curie le 11 décembre 2012.
- Conférence d’1 heure 30 (titre : « Annotation des ARN non codants chez les Eucaryotes — les enjeux biologiques ») à la journée *DE^CODAGE* du GdR BiM le 26 avril 2013.
- Conférence de 45 minutes (titre : « Statistical analysis of small RNA-seq data ») au symposium *Statistical analysis of RNA-seq data : advances and challenges* de l’Institut Pasteur le 26 novembre 2013.

(donc en moyenne : environ 6h30 d’interventions par an)

3.3 Participation à l’organisation de conférences

(aucune)

3.4 Participation à des revues ou ouvrages de vulgarisation

Un article dans une revue de vulgarisation :

Pinzón Restrepo, N., Martinez, L., Seitz, H. et Cavallé, J. (2013) Les petits ARN entrent dans l’arène. *Pour la science*, dossier n°81 « L’hérédité sans gènes » (octobre-décembre 2013). Sommaire du numéro.

Interventions orales :

- Cours de deux heures, sur le thème « Petits ARN régulateurs chez les Eucaryotes » le 9 juin 2010, à destination d’enseignants en SVT dans le secondaire (formation organisée par l’IRFEC, à Toulouse).
- Intervention de 30 minutes, sur les petits ARN régulateurs chez les Animaux, le 19 novembre 2010, au Carrefour Eurobiomed (journée de rencontre entre industriels et chercheurs académiques), à Montpellier.
- Interventions sur le thème de « la cellule » auprès de scolaires (quatre groupes d’élèves de terminale, un groupe d’élèves de 6ème et de CE2), organisées par l’association « Connaisciences », 31 mars et 1^{er} avril 2011.
- Fête de la science 2013 : intervention de 30 minutes sur le thème « Qu’est-ce que la génétique ? » les 10 et 11 octobre 2013 (devant des groupes de scolaires) et le 12 octobre 2013 (devant le grand public).

3.5 Interventions dans la presse écrite et audiovisuelle

Mai 2013 : conseil scientifique auprès de Damien Durand, journaliste du site <http://www.atlantico.fr>.

3.6 Participation à des travaux d'expertise

- Rapporteur pour les journaux *Biochimie*, *BioEssays*, *Cell Reports*, *Cell Research*, *Current Biology*, *EMBO Reports*, *Genome Biology*, *Genome Research*, *Genomics*, *Molecular Oncology*, *Nucleic Acids Research*, *PLoS One*, *RNA*, *RNA Biology* et *Silence*.
- Rapporteur pour l'appel d'offre « Deep-sequencing » du GIS IBiSA (Génoscope et Centre national de séquençage) (2010).
- Rapporteur pour l'appel d'offre de la FRM de la région Alsace (2011).
- Rapporteur pour le programme « ANR Blanche » de l'ANR, panel SVSE2 (2012).
- Membre du comité de programme du congrès JOBIM (Journées Ouvertes en Biologie, Informatique et Mathématiques) 2012 et 2013.
- Rapporteur pour les thèses de Caroline Jacquier (directeur de thèse : Christophe Antoniewski, Institut Pasteur, Paris; 28 juin 2010), de Chongjian Chen (directeurs de thèse : Daniel Gautheret, université Paris Sud, Orsay, et Liang-Hu Qu, Sun Yat-Sen University, Canton, Chine; 17 décembre 2010), d'Antoine Molaro (directeurs de thèse : Phillip Avner, Institut Pasteur, Paris, et Gregory Hannon, Cold Spring Harbor Laboratory, NY, États-Unis; 9 mai 2012), d'Ilyass Zniber (directeur de thèse : Denis Dupuy, IECB, Bordeaux; 17 décembre 2012) et de Lue Huang (directeur de thèse : François Dautry, ENS Cachan; 19 décembre 2012).
- Examineur pour les thèses de Nicolas Greliche (directeur de thèse : David-Alexandre Trégoët; 18 février 2013) et de Thomas Grentzinger (directrice de thèse : Séverine Chambeyron; 13 juin 2013).
- Rapporteur pour l'HDR de Séverine Chambeyron (directeur des recherches : Alain Bucheton, IGH, Montpellier; 11 mai 2011).
- Membre du jury de recrutement d'un maître de conférence pour l'université Toulouse III Paul Sabatier (poste n°64MCF0300) en avril et mai 2013.

4 Transfert technologique, relations industrielles et valorisation

J'ai participé, au cours de cette période, à trois contrats de recherche :

- *Career Development Award* de Human Frontier Science Program (CDA n°00017-2010-C) : d'août 2010 à août 2013. Rôle : coordonnateur du projet. Partenaire : aucun. Montant : 100 000 \$ par an, pendant les trois ans. Travaux effectués : mesure de la corrélation entre abondance des ARNm et conservation de leurs sites de complémentarité aux miRNA, qui suggère que les ARNm endogènes jouent le rôle de titrateurs de miRNA.
- ATIP-Avenir du CNRS et de l'Inserm, avec financement complémentaire de la part de Sanofi (partenaire du programme ATIP-Avenir) : de mars 2012 à février 2015. Rôle : coordonnateur du projet. Partenaire : aucun. Montant : 120 000 € par an, pendant les trois ans. Travaux effectués : mesure de la robustesse des cibles prédites de miR-223 dans les neutrophiles de Souris, vis-à-vis de la répression guidée par ce miRNA. Comparaison de l'abondance de miRNA et de leurs cibles prédites au cours de la différenciation des cellules C2C12. Ces résultats suggèrent que la fonction principale des sites de complémentarité conservés aux miRNA est la modulation des miRNA par titration, et non la répression des ARNm qui les portent.
- Projet ANR « miRNAadapt » (coordonnateur : Denis Tagu, INRA). Rôle : partenaire. Montant : 21 000 €. Mes travaux consistent à participer aux prédictions de cibles de miRNA chez le Puceron *Acyrtosiphon pisum*, et à développer des critères de sélection des cibles les plus probablement fonctionnellement ciblées par les miRNA.

J'ai également effectué une mission de consultance (3 mois : mars à mai 2012) auprès de l'entreprise « Medesis Pharma » (c'est une start-up qui développe des vecteurs pour l'adressage de molécules actives *in vivo*, qui aimerait utiliser ses produits pour l'administration de siRNA et qui cherchait de l'aide pour choisir les séquences de siRNA les plus efficaces). Temps passé : 15 heures.

5 Encadrement, animation et management de la recherche

Je dirige depuis octobre 2011 une équipe de recherche à l'IGH (UPR 1142 du CNRS). L'évolution de l'effectif de l'équipe est schématisée en figure 5 : l'équipe actuelle compte 5 membres (une assistante ingénieure, deux post-doctorants, une chargée de recherche, et moi-même).

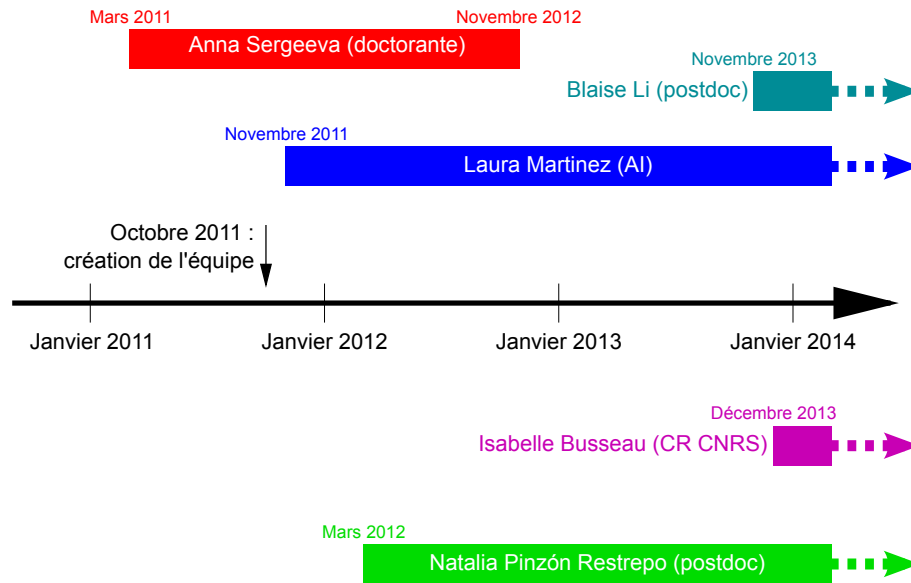


FIGURE 5 – Effectifs de l'équipe.

6 Objectifs / Projet de recherche

Nos projets pour les prochaines années viseront essentiellement à explorer les implications biologiques de l'hypothèse des pseudo-cibles, qui a été confortée par nos résultats de ces dernières années (voir section 2.1, page 6).

Nous allons utiliser une approche similaire à celle de l'expérience de mesure de la robustesse de l'activité des cibles prédites de miR-223 dans les neutrophiles de Souris, vis-à-vis des variations de leur niveau d'expression (voir figure 2, page 10), chez la Drosophile. Cette expérience nous permettra d'identifier les gènes dont l'expression est la plus homogène entre les individus, parmi les cibles prédites pour quelques miRNA : ces gènes seront les candidats les plus crédibles pour être sensibles au dosage (et donc, pour être de véritables cibles de miRNA, au sens fonctionnel). Chez la Drosophile, il nous sera ensuite possible de muter les sites de fixation du miRNA sur leurs ARNm, et d'évaluer la contribution de ces cibles au phénotype d'un mutant du miRNA, en d'autres termes : de vérifier le rôle biologique de ces interactions. Cette expérience constituerait la première identification, à grande échelle, des cibles de miRNA qui contribuent à un phénotype macroscopique.

Dans le cadre de ce projet, nous nous intéresserons aux miRNA miR-124, miR-3, miR-4 et miR-5, qui sont, chacun, exprimés principalement dans l'embryon (Aravin *et al.*, 2003 ; Sempere *et al.*, 2003 ; Sun *et al.*, 2012). Il faudra tout d'abord mesurer l'amplitude de la répression de ces miRNA sur leurs cibles prédites, dans les cellules qui expriment le miRNA. Nous utiliserons, pour cela, la méthode décrite par Sun *et al.*, 2012, qui consiste à faire exprimer un marqueur fluorescent (sous le contrôle du promoteur du miRNA) par les cellules qui expriment le miRNA, et à trier les cellules des embryons dissociés, par cytométrie de flux. Nous réaliserons cette expérience à la fois sur quelques réplicats de lots d'embryons sauvages, et d'embryons mutants pour le miRNA, de manière à mesurer l'effet du miRNA sur chacun de ces gènes *in vivo* (Sun *et al.*, 2012). Strictement parlant, il faudrait mesurer l'expression de ces gènes par l'abondance de leurs protéines, mais la protéomique quantitative n'est pas aussi sensible que la transcriptomique quantitative (voir Selbach *et al.*, 2008 ; Baek *et al.*, 2008 par exemple) et les variations d'abondance de protéines semblent être fidèlement représentées par les variations d'abondance de leurs ARNm en régime stationnaire (Guo *et al.*, 2010 ; Bazzini *et al.*, 2012). Ce sont donc les ARNm des cibles prédites, que nous quantifierons, par RNA-seq. Nous mesurerons également la variabilité inter-individus dans l'expression de ces gènes, en utilisant la même procédure, mais sur des embryons sauvages individuels. La quantité de cellules purifiées dans chacune de ces expériences pourrait être faible, mais des protocoles spéciaux de RNA-seq ont été développés pour analyser de petits échantillons (jusqu'à la cellule unique ; Hashimshony *et al.*, 2012 ; Ramsköld *et al.*, 2012). Les mutants de miRNA qui ne sont pas encore disponibles dans la communauté seront générés selon la procédure décrite par Sun *et al.*, 2012.

Dans l'expérience que nous avons déjà réalisée, dans les neutrophiles de Souris, seuls 10% des gènes étaient exprimés à des niveaux relativement constants entre individus (*i.e.* : les variations inter-individus étaient moins grandes que l'effet du miRNA sur le gène). Ce nombre de gènes est suffisamment petit pour permettre une étude détaillée, gène par gène, des relations d'épistasie entre le miRNA et l'ARNm. Si le nombre de gènes à expression peu variable est du même ordre de grandeur chez la Drosophile que chez la Souris, nous pourrions donc vérifier, gène par gène, la contribution de ces candidats aux phénotypes des mutants de ces miRNA.

Nous allons également rechercher, chez la Drosophile, d'éventuels titrateurs de miRNA qui ne seraient pas, simultanément, des ARNm. Ces ARN non-codants auraient pour seule fonction la régulation des miRNA (ce qui faciliterait l'interprétation du phénotype éventuel dû à leur mutation). Une analyse informatique du génome non-codant de la Drosophile m'a permis d'identifier 32 loci portant des sites de complémentarité au miRNA *bantam* (dont le phénotype est facile à évaluer *in vivo* : à l'état hétérozygote, le mutant *bantam* présente des défauts de croissance générale, et de morphologie et de taille des ailes : Hipfner *et al.*, 2002 ; Brennecke *et al.*, 2003). Nous savons déjà qu'au moins 5 de ces loci sont transcrits, et génèrent donc des ARN non-codants, qui n'avaient pas encore

été décrits, et qui semblent pouvoir interagir avec le miRNA (voir section 2.1). Nous allons tester l'expression de la totalité des 32 candidats, par RT-PCR sur les différents stades développementaux de la Drosophile. Afin d'évaluer l'effet de titration que ces ARN pourraient exercer sur le miRNA, nous allons mesurer la croissance et estimer la morphologie des ailes sur des mutants de ces gènes non-codants (si des altérations génétiques de ces loci sont disponibles dans la communauté) ou en les réprimant par RNAi transgénique. Nous évaluerons la croissance de ces mouches selon le protocole rigoureux mis en place par Hipfner *et al.*, 2002. Il nous faudra d'abord identifier les extrémités 5' et 3' de ces ARN non-codants, pour pouvoir identifier les mutations disponibles dans ces gènes, ou préparer des transgènes ciblant ces ARN par RNAi.

L'hypothèse des pseudo-cibles implique également que l'identité précise de l'ARNm qui porte les sites de fixation qui titrent le miRNA n'a pas d'importance en soi, et une pseudo-cible pourrait être remplacée, dans son rôle de modulateur de miRNA, par d'autres ARNm de patron d'expression similaire. De fait, sur de longues distances évolutives, il est prévisible que la liste des pseudo-cibles varie. Les cibles réelles, en revanche, devraient porter des sites de fixation du miRNA plus conservés : si elles perdent leur interaction avec le miRNA, la fonction du miRNA lui-même est perdue. Si les cibles réelles portent des sites de fixation au miRNA mieux conservés que ceux des pseudo-cibles, alors les gènes les plus dose-sensibles devraient porter les sites de fixation de miRNA les mieux conservés (puisque c'est la sensibilité de la fonction vis-à-vis du niveau d'expression qui distingue les cibles réelles des pseudo-cibles).

Quelques miRNA sont très conservés parmi les Bilatériens, et notamment *let-7*, qui contrôle des phénotypes faciles à évaluer (parmi les 5 miRNA pan-bilatériens, *let-7* est le seul qui contrôle la viabilité : Meneely et Herman, 1979 ; Sokol *et al.*, 2008). Nous allons donc chercher les gènes dont l'interaction avec *let-7* est la mieux conservée entre les Mammifères, la Drosophile et le Nématode (d'après Chen et Rajewsky, 2006, trois gènes ont un site de fixation conservé à *let-7* entre *Homo sapiens* et *Drosophila melanogaster*, et quatre gènes ont un site conservé entre *H. sapiens* et *Caenorhabditis elegans*). L'analyse effectuée par Chen et Rajewsky, 2006 est probablement trop stringente, puisque le gène *lin-41* (chez *C. elegans*) / *Trim71* (chez les Mammifères) est absent de la liste, alors que sa 3' UTR contient un site conservé entre ces deux lignages. Nous commencerons donc par refaire cette analyse informatique, avec des paramètres mieux optimisés.

Nous évaluerons ensuite la sensibilité de l'activité de chacun de ces gènes vis-à-vis de la répression guidée par *let-7*, en délétant leurs sites d'interaction à *let-7* chez la Drosophile par la technique CRISPR. Nous mesurerons la viabilité et la fertilité de ces mutants, ainsi que celles d'une cohorte de même taille, de mutants sur des gènes contrôles (délétés pour les sites de fixation à *let-7* peu conservés : ces sites seront conservés parmi les Drosophilides, mais pas chez le Nématode ni les Mammifères). Cette cohorte de contrôle inclura notamment les gènes *hairy*, *ribbon* et *yantar*, qui contrôlent des phénotypes développementaux très bien caractérisés (nous mesurerons également ces phénotypes, en plus de la viabilité et de la fertilité). Un doctorant sera recruté pour mener ce projet, à la fois sur ses aspects bio-informatiques et sur ses aspects expérimentaux. Il bénéficiera pour ces deux disciplines de l'expertise d'Isabelle Busseau (chargée de recherche nouvellement arrivée dans l'équipe) et de moi-même.

Chacun de ces trois projets implique à la fois des expériences de biologie moléculaire à haut débit, des analyses informatiques à l'échelle du génome, ou de la collection de génomes, et des explorations fonctionnelles de phénotypes développementaux chez la Drosophile. Ces trois méthodologies font partie du cœur de métier de notre unité, l'IGH. Les analyses génomiques et la régulation du développement par les petits ARN font partie des priorités reconnues de l'unité, et les techniques qu'elles mettent en œuvre sont également utilisées en routine par de nombreuses équipes de l'institut. De cette manière, nous sommes sûrs de pouvoir bénéficier de l'expérience de nos collègues, et réciproquement, l'expertise que nous acquièrerons à l'occasion de ces travaux sera certainement profitable à tous.

Enfin, à plus long terme, l'hypothèse des pseudo-cibles pourrait ouvrir une nouvelle perspective sur notre compréhension de la régulation des gènes. Une cible régulatrice n'est pas seulement un gène qui est affecté par une voie de régulation : c'est un gène qui est affecté *suffisamment*. L'amplitude de la régulation doit être confrontée à la robustesse du système biologique vis-à-vis des fluctuations. Ce concept peut être facilement généralisé aux autres types de régulateurs, outre les miRNA : facteurs de transcription, protéines d'interaction à l'ARN, modificateurs post-traductionnels, *etc.* Depuis plusieurs décennies, des expériences de biologie moléculaire ont identifié des relations entre régulateurs et cibles, qui ont un effet mesurable sur la cible à l'échelle moléculaire. Mais une proportion de ces « cibles » pourrait ne pas être régulée assez pour que ces événements microscopiques se traduisent en effets physiologiques. Si ces interactions ont une fonction (par exemple, si elles ont été sélectivement conservées pendant l'évolution), alors cette fonction pourrait être de réguler le régulateur, par titration.

La généralisation de ce concept aux facteurs de transcription pourrait expliquer des observations intrigantes sur le nombre de cibles identifiées, et la faible conservation de nombreux sites de fixation de facteurs de transcription. Des sites de fixation déterminés expérimentalement pour plusieurs facteurs de transcription sont, curieusement, peu conservés entre les Vertébrés, ce qui suggère qu'une grande majorité des sites de fixation de facteurs de transcription sont neutres du point de vue de la sélection naturelle (Odom *et al.*, 2007 ; Schmidt *et al.*, 2010). Cette observation surprenante pourrait signifier que l'activité physiologique de la plupart des cibles est insensible aux changements d'expression induits par ces facteurs de transcription. De manière similaire, la généralisation de l'hypothèse des pseudo-cibles pourrait expliquer la faible spécificité apparente des protéines d'interaction à l'ARN : l'identification à haut débit des partenaires ARN de quelques-unes de ces protéines a révélé des milliers de « cibles » ARN pour chacune (Hafner *et al.*, 2010 ; Lebedeva *et al.*, 2011). Si une grande partie du transcriptome est affectée par ces protéines, la spécificité de ces régulations ne peut pas venir de la spécificité de l'interaction elle-même, elle découlerait donc de la sensibilité différente de tous ces gènes vis-à-vis de ce phénomène moléculaire.

De ce point de vue, ni la description moléculaire des phénomènes biologiques, ni les analyses phylogénétiques, ne peuvent expliquer la nature réelle de la relation entre les deux interacteurs. La biologie moléculaire peut identifier des interactions moléculaires ; la génomique comparative peut montrer qu'elles ont une fonction biologique. Mais aucune des deux ne peut déterminer si la fonction consiste à réguler la cible, ou le régulateur, par titration. Seule la biologie des systèmes peut aborder cette question, en quantifiant les conséquences de cette interaction sur l'activité biologique de chaque composant, cible et régulateur.

Ces questions, qui vont probablement prendre une place importante dans mes recherches des prochaines années, sont au cœur des priorités annoncées de deux domaines de recherche de l'Institut des sciences biologiques du CNRS : les domaines « Biologie intégrative animale et végétale » (<http://www.cnrs.fr/insb/presentation/domaines/biologieintegrative.htm>) et « Génétique, génomique et expression des gènes » (<http://www.cnrs.fr/insb/presentation/domaines/gene.htm>).

7 Références

- Aravin, A. A., Lagos-Quintana, M., Yalcin, A., Zavolan, M., Marks, D., Snyder, B., Gaasterland, T., Meyer, J., et Tuschl, T. (2003). The small RNA profile during *Drosophila melanogaster* development. *Dev Cell*, 5(2) : 337–350.
- Aravin, A. A., Sachidanandam, R., Girard, A., Fejes-Toth, K., et Hannon, G. J. (2007). Developmentally regulated piRNA clusters implicate MILI in transposon control. *Science*, 316(5825) : 744–747.
- 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.
- Bazzini, A. A., Lee, M. T., et Giraldez, A. J. (2012). Ribosome profiling shows that miR-430 reduces translation before causing mRNA decay in zebrafish. *Science*, 336(6078) : 233–237.
- Brennecke, J., Aravin, A. A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., et Hannon, G. J. (2007). Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell*, 128(6) : 1089–1103.
- 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.
- Chekulaeva, M. et Filipowicz, W. (2009). Mechanisms of miRNA-mediated post-transcriptional regulation in animal cells. *Curr Opin Cell Biol*, 21(3) : 452–460.
- Chen, J. F., Mandel, E. M., Thomson, J. M., Wu, Q., Callis, T. E., Hammond, S. M., Conlon, F. L., et Wang, D. Z. (2006). The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. *Nat Genet*, 38(2) : 228–233.
- Chen, K. et Rajewsky, N. (2006). Deep conservation of microRNA-target relationships and 3' UTR motifs in vertebrates, flies, and nematodes. *Cold Spring Harb Symp Quant Biol*, 71 : 149–156.
- Chung, W. J., Okamura, K., Martin, R., et Lai, E. C. (2008). Endogenous RNA interference provides a somatic defense against *Drosophila* transposons. *Curr Biol*, 18(11) : 795–802.
- 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.
- 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, M., Seitz, H., Horwich, M. D., Li, C., Du, T., Lee, S., Xu, J., Kittler, E. L., Zapp, M. L., Weng, Z., et Zamore, P. D. (2008). Endogenous siRNAs derived from transposons and mRNAs in *Drosophila* somatic cells. *Science*, 320(5879) : 1077–1081.
- Grimson, A., Srivastava, M., Fahey, B., Woodcroft, B. J., Chiang, H. R., King, N., Degan, B. M., Rokhsar, D. S., et Bartel, D. P. (2008). Early origins and evolution of microRNAs and Piwi-interacting RNAs in animals. *Nature*, 455(7217) : 1193–1197.
- Guo, H., Ingolia, N. T., Weissman, J. S., et Bartel, D. P. (2010). Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature*, 466(7308) : 835–840.
- Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, Jr., M., Jungkamp, A. C., Munschauer, M., Ulrich, A., Wardle, G. S., Dewell, S., Zavolan, M., et Tuschl, T. (2010). Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell*, 141(1) : 129–141.
- Hashimshony, T., Wagner, F., Sher, N., et Yanai, I. (2012). CEL-Seq : single-cell RNA-Seq by multiplexed linear amplification. *Cell Rep*, 2(3) : 666–673.
- Hipfner, D. R., Weigmann, K., et Cohen, S. M. (2002). The *bantam* gene regulates *Drosophila* growth. *Genetics*, 161(4) : 1527–1537.

- Kawamata, T., Seitz, H., et Tomari, Y. (2009). Structural requirement of miRNAs for RISC-loading and Slicer-independent unwinding. *Nat Struct Mol Biol*, 16(9) : 953–960.
- Kawamura, Y., Saito, K., Kin, T., Ono, Y., Asai, K., Sunohara, T., Okada, T. N., Siomi, M. C., et Siomi, H. (2008). *Drosophila* endogenous small RNAs bind to Argonaute 2 in somatic cells. *Nature*, 453(7196) : 793–797.
- Kim, H. K., Lee, Y. S., Sivaprasad, U., Malhotra, A., et Dutta, A. (2006). Muscle-specific microRNA miR-206 promotes muscle differentiation. *J Cell Biol*, 174(5) : 677–687.
- Lebedeva, S., Jens, M., Theil, K., Schwanhäusser, B., Selbach, M., Landthaler, M., et Rajewsky, N. (2011). Transcriptome-wide analysis of regulatory interactions of the RNA-binding protein HuR. *Mol Cell*, 43(3) : 340–352.
- Meneely, P. M. et Herman, R. K. (1979). Lethals, steriles and deficiencies in a region of the X chromosome of *Caenorhabditis elegans*. *Genetics*, 92(1) : 99–115.
- Odom, D. T., Dowell, R. D., Jacobsen, E. S., Gordon, W., Danford, T. W., MacIsaac, K. D., Rolfe, P. A., Conboy, C. M., Gifford, D. K., et Fraenkel, E. (2007). Tissue-specific transcriptional regulation has diverged significantly between human and mouse. *Nat Genet*, 39(6) : 730–732.
- Okamura, K., Chung, W. J., Ruby, J. G., Guo, H., Bartel, D. P., et Lai, E. C. (2008). The *Drosophila* hairpin RNA pathway generates endogenous short interfering RNAs. *Nature*, 453(7196) : 803–806.
- Rajasethupathy, P., Antonov, I., Sheridan, R., Frey, S., Sander, C., Tuschl, T., et Kandel, E. R. (2012). A role for neuronal piRNAs in the epigenetic control of memory-related synaptic plasticity. *Cell*, 149(3) : 693–707.
- Ramsköld, D., Luo, S., Wang, Y. C., Li, R., Deng, Q., Faridani, O. R., Daniels, G. A., Khrebtkova, I., Loring, J. F., Laurent, L. C., Schroth, G. P., et Sandberg, R. (2012). Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat Biotechnol*, 30(8) : 777–782.
- Schmidt, D., Wilson, M. D., Ballester, B., Schwalie, P. C., Brown, G. D., Marshall, A., Kutter, C., Watt, S., Martinez-Jimenez, C. P., Mackay, S., Talianidis, I., Flicek, P., et Odom, D. T. (2010). Five-vertebrate ChIP-seq reveals the evolutionary dynamics of transcription factor binding. *Science*, 328(5981) : 1036–1040.
- Seitz, H. (2009). Redefining microRNA targets. *Curr Biol*, 19(10) : 870–873.
- Seitz, H., Tushir, J. S., et Zamore, P. D. (2011). A 5'-uridine amplifies miRNA/miRNA* asymmetry in *Drosophila* by promoting RNA-induced silencing complex formation. *Silence*, 2 : 4.
- 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.
- Sempere, L. F., Sokol, N. S., Dubrovsky, E. B., Berger, E. M., et Ambros, V. (2003). Temporal regulation of microRNA expression in *Drosophila melanogaster* mediated by hormonal signals and broad-Complex gene activity. *Dev Biol*, 259(1) : 9–18.
- Sokol, N. S., Xu, P., Jan, Y. N., et Ambros, V. (2008). *Drosophila let-7* microRNA is required for remodeling of the neuromusculature during metamorphosis. *Genes Dev*, 22(12) : 1591–1596.
- Sun, K., Westholm, J. O., Tsurudome, K., Hagen, J. W., Lu, Y., Kohwi, M., Betel, D., Gao, F. B., Haghghi, A. P., Doe, C. Q., et Lai, E. C. (2012). Neurophysiological defects and neuronal gene deregulation in *Drosophila* mir-124 mutants. *PLoS Genet*, 8(2) : e1002515.
- Tam, O. H., Aravin, A. A., Stein, P., Girard, A., Murchison, E. P., Cheloufi, S., Hodges, E., Anger, M., Sachidanandam, R., Schultz, R. M., et Hannon, G. J. (2008). Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature*, 453(7194) : 534–538.
- Tsutsumi, A., Kawamata, T., Izumi, N., Seitz, H., et Tomari, Y. (2011). Recognition of the pre-miRNA structure by *Drosophila* Dicer-1. *Nat Struct Mol Biol*. In press.

- Watanabe, T., Totoki, Y., Toyoda, A., Kaneda, M., Kuramochi-Miyagawa, S., Obata, Y., Chiba, H., Kohara, Y., Kono, T., Nakano, T., Surani, M. A., Sakaki, Y., et Sasaki, H. (2008). Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature*, 453(7194) : 539–543.
- 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.