#### RESUME

Du fait de la rotation de la terre sur son axe, les organismes vivants se trouvent confrontés à un changement quotidien de leur environnement correspondant à l'alternance des jours et des nuits. Parce qu'il y a un avantage sélectif à anticiper ces variations journalières, la majorité des espèces ont développé des horloges endogènes qui leur permettent cette anticipation. Chez les mammifères, cette horloge endogène contrôle la rythmicité de nombreuses fonctions qui vont osciller avec une période proche de 24h, dite circadienne. La rythmicité d'un grand nombre de ces fonctions, cellulaires, physiologiques ou comportementales, repose sur une expression rythmique de gènes et dans chaque tissu, le transcriptome circadien est estimé entre 8 à 40% des gènes exprimés. Le mécanisme moléculaire à la base du fonctionnement des oscillateurs circadiens implique des boucles de rétrocontrôle de la transcription, et ce mécanisme transcriptionnel permet de contrôler la rythmicité d'expression de nombreux gènes cibles. Cependant plusieurs études récentes montrent que 80% des gènes rythmiques sont contrôlés par des mécanismes post-transcriptionnels.

Au cours de notre étude nous avons identifié un mécanisme post-transcriptionnel impliqué dans le contrôle de l'expression rythmique de gènes. Ce mécanisme implique des corps nucléaires appelés paraspeckles, présents dans les cellules de mammifères. Les paraspeckles sont organisés autour d'un long ARN non codant (lncRNA), Neat1, auquel s'associent plusieurs protéines de liaison aux ARNs. Fonctionnellement ces paraspeckles peuvent retenir dans le noyau des ARNs, particulièrement les ARNm possédant dans leur région 3' non codante (3'UTR) des motifs de type IRAlu.

Nous avons montré dans une lignée de cellules hypophysaires somatolactotropes (GH4C1) que les composants principaux des paraspeckles, dont Neat1, sont exprimés de manière circadienne, ce qui induit une variation circadienne du nombre de paraspeckles dans le noyau des cellules. En introduisant une séquence de type IRAlu dans la région 3'UTR du gène rapporteur Egfp et en faisant exprimer de manière stable cette construction par des cellules GH4C1, nous avons montré un rythme circadien de la rétention nucléaire de l'ARNm Egfp associé à un rythme circadien de la protéine EGFP détectée par vidéo-microscopie en temps réel. Lorsque les paraspeckles sont détruits après transfection des cellules par des siRNA ou des oligonucléotides antisens dirigés contre Neat1, ces rythmes disparaissent.

Pour connaître l'ensemble des ARNm endogènes qui peuvent être retenus par les paraspeckles, nous avons développé une technique de « RNA pull-down » qui permet d'isoler les partenaires moléculaires d'un long ARN non codant. En couplant cette approche à du séquençage haut débit des ARNs, nous avons établi la liste des ARNs associés à Neat1 et donc retenus par les paraspeckles. L'analyse des régions 3'UTR de ces ARNs n'a pas permis d'identifier des motifs IRSINE, analogues chez le rat des motifs IRAlus des primates.

Dans le but d'identifier d'autres déterminants moléculaires présents dans les ARNm capables de s'associer aux paraspeckles, nous avons sélectionné trois ARNm cibles de Neat1, et nous avons inséré leurs régions 3'-UTR, intégrales ou bien fragmentées, en aval de la partie codante du gène rapporteur Egfp. Cette étude nous a permis de conclure que la reconnaissance par les paraspeckles peut impliquer des régions autres que la région 3'UTR des ARNs, et que dans la région 3'UTR, plusieurs motifs agissant de concert participent à la liaison aux paraspeckles.

En conclusion, l'ensemble des travaux réalisés nous a permis d'identifier un nouveau mécanisme post-transcriptionnel permettant l'expression circadienne de gènes dans les cellules hypophysaires, mais aussi potentiellement dans d'autres structures de l'organisme et d'autre part d'attribuer un rôle physiologique dans le fonctionnement du système circadien à un long ARN non-codant et à des sous-domaines nucléaires dont les fonctions demeurent énigmatiques.

### **CURRICULUM VITAE**

### **Manon TORRES**

Née le 11 Juillet 1991 +33 6 24 75 81 29 torres.manon@gmail.com

### Scolarité

2014 – présent	Doctorante contractuelle en Neurosciences, Aix-Marseille Université, Marseille,
	France
2012-2014	Master en Neurosciences, Aix-Marseille Université, Marseille, France
2009-2012	Licence en Physiologie animale et Neurosciences, Université de Montpellier-2,
	Montpellier, France

### **Expérience professionnelle**

2014 – présent	Doctorat, Centre de recherche en Neurobiologie et Neurophysiologie de Marseille,
	CNRS- UMR7286, Marseille, France
2013-2014	Stage de Master 2, Centre de recherche en Neurobiologie et Neurophysiologie de
	Marseille, CNRS- UMR7286, Marseille, France
2013	Stage de Master 1, Laboratoire de Physiologie et Physiopathologie du système
	nerveux Somatomoteur, Marseille, France

### **Publications**

M Torres and A Kramer

Mealtime is NONO-speckled: timing hepatic adaptation to food Sous presse, **Cell Metabolism** 

<u>M Torres</u>, D Becquet, S Guillen, B Boyer, M Moreno, MP Blanchard, JL Franc & AM François-Bellan A RNA pull-down procedure to identify RNA targets of a long noncoding RNA Sous presse, **JoVE** 

<u>M Torres</u>, D Becquet, MP Blanchard, S Guillen, B Boyer, M Moreno, JL Franc & AM François-Bellan *Circadian processes in RNA's life cycle* Sous presse, **WIREs RNA** 

<u>M Torres</u>, D Becquet, MP Blanchard, S Guillen, B Boyer, M Moreno, JL Franc & AM François-Bellan *Paraspeckles as rhythmic nuclear mRNA anchorages responsible for circadian gene expression*, **Nucleus**, 2017 DOI: 10.1080/19491034.2016.1277304

<u>M Torres</u>, D Becquet, MP Blanchard, S Guillen, B Boyer, M Moreno, JL Franc & AM François-Bellan *Circadian RNA expression elicited by 3'-UTR IRAlu-paraspeckle associated elements* **eLife**. 2016 DOI:10.7554/eLife.14837.

### **Présentations orales et posters**

### 2017 – 42<sup>ème</sup> congrès de la Société de Neuroendocrinologie

Présentation orale - « RNA pull-down » a new technology to identify RNA targets of a long noncoding RNA

### 2017 – XV European Biological Rhythms Society (EBRS)

Poster – Paraspeckles as nuclear bodies driving non-transcriptional circadian mRNA expression

### 2016 – 41<sup>ème</sup> congrès de la Société de Neuroendocrinologie

Présentation orale – mRNA circadian rhythmicity depend of IRAlu elements recognized by paraspeckles.

### 2016 – 45<sup>ème</sup> congrès de la Société Francophone de chronobiologie

Présentation orale – Pituitary transcriptoma rhythicity depend of IRAlu elements recognized by paraspeckles.

### **2015 – XIV European Biological Rhythms Society (EBRS) and IV World Congress of Chronobiology (WCC)** Poster - *Neat1 a long non-coding RNA involved in pituitary circadian rhythms : characterization of its interactome.*

#### **2015 – 5<sup>ème</sup> congrès de la Société Française d'Endocrinologie et Diabétologie Pédiatrique** Oral communication – Circadian rhythmicity is altered in Prader-Willi syndrome

#### **Bourses et prix**

2017 – 42<sup>ème</sup> congrès de la Société de Neuroendocrinologie – Bourse de voyage
2017 – XV European Biological Rhythms Society (EBRS) – Bourse de voyage
2016 – 41<sup>ème</sup> congrès de la Société de Neuroendocrinologie – Prix pour communication orale
2016 – 41<sup>ème</sup> congrès de la Société de Neuroendocrinologie – Bourse de voyage
2016 – 45<sup>ème</sup> congrès de la Société Francophone de chronobiologie – Prix pour communication orale
2016 – 45<sup>ème</sup> congrès de la Société Francophone de chronobiologie – Bourse de voyage
2016 – 45<sup>ème</sup> congrès de la Société Francophone de chronobiologie – Bourse de voyage
2016 – 45<sup>ème</sup> congrès de la Société Francophone de chronobiologie – Bourse de voyage
2013 – Société Française d'Endocrinologie et Diabétologie Pédiatrique (SFEDP) – Bourse Merck-Serono

### Autres

**2015 – Semaine du cerveau « le ventre notre deuxième cerveau »** - Réalisation du diaporama et présentation dans des lycées.

**2014 – Diplôme d'expérimentation animale** - Connaissance et protection de l'animal de laboratoire, Méthodologie expérimentale, niveau 1.

# AIX-MARSEILLE UNIVERSITÉ

Faculté de Médecine de Marseille

Ecole doctorale des Sciences de la Vie et de la Santé

# THESE

Spécialité : Neurosciences

Par

**Manon Torres** 

# Rôle du long ARN non codant Neat1 dans la rythmicité circadienne hypophysaire

Présentée et publiquement soutenue le 26 janvier 2018 devant la commission d'examen composée de :

Dr. Valérie Simonneaux	Rapporteur
Dr. Frédéric Gachon	Rapporteur
Dr. Michel Khrestchatisky	Examinateur
Dr. Anne-Marie François-Bellan	Directrice de thèse

Centre de Recherche en Neurobiologie-Neurophysiologie de Marseille UMR 7286, faculté de Médecine secteur Nord Bd Pierre Dramard, 13344 Marseille

Introduction générale	3
Revue de littérature	6
I – Les rythmes circadiens	7
1. Les rythmes biologiques	7
1.1. Importance des rythmes biologiques	7
1.2. Les rythmes circadiens	8
1.2.1. Découverte des rythmes circadiens	
1.2.2. Définition des rythmes circadiens	9
2. Organisation du système circadien	11
2.1. L'horloge centrale chez les mammifères	
2.1.1. Mise en évidence de l'horloge centrale dans le noyau suprachiasmatique	
2.1.2. Anatomie et organisation du noyau suprachiasmatique	12
2.1.3. Fonctionnement moléculaire de l'oscillateur circadien	13
2.2. Synchronisation externe du système circadien	17
2.2.1 Voies anatomiques d'entrée sur le NSC	17
2.2.2. Synchronisation photique	
2.2.3. Synchronisation non-photique	20
2.3. Les oscillateurs périphériques	21
2.3.1. Découverte et démonstration de l'existence des oscillateurs périphériques	21
2.3.2. Robustesse des oscillateurs périphériques	
2.4. Les mécanismes de synchronisation interne du système circadien	23
2.4.1 Voies anatomiques de sortie du NSC	
2.4.2 Synchronisation des oscillateurs périphériques	
3. L'hypophyse	
3.1. Structure et fonctions de l'hypophyse :	
3.2. Rythmicité de l'hypophyse	27
3.3. L'hypophyse : un oscillateur autonome	28
3.4. La lignée cellulaire hypophysaire GH4C1	28
4. Les rythmes d'expression génique	29
4.1. Importance quantitative des transcriptomes cellulaires	29
4.2. Mécanismes transcriptionnels de contrôle	30
4.3. Mécanismes post-trancriptionnels	
II - Le long ARN non codant Neat1 et les paraspeckles	61
1. Les longs ARN non codants	61
1.1. Définition générale des longs ARN non codants	61
1.2. Les fonctions des longs ARN non codants	63
2. Les paraspeckles	65
2.1. Découverte des paraspeckles	65
2.2. Les éléments du paraspeckles	65
2.2.1. Le long ARN non codant Neat1	65
2.2.2. Les protéines	67
2.2.3. Assemblage des paraspeckles	68
2.3. Fonctions des paraspeckles	70
2.3.1. Rétention nucléaire des protéines	
2.3.2. Rétentions d'ARNm	71
3. Les fonctions physiologiques et pathologiques associées à Neat1 et aux	
paraspeckles	74
3.1. Fonctions physiologiques de Neat1	74
3.2. Les pathologies associées à Neat1	77

Résultats expérimentaux	79
A RNA pull down procedure to indentify RNA targets of a long noncoding RNA	80
Introduction	81
Résultats	82
Conclusion	83
Circadian RNA expression elicited by 3'-UTR IRAlu-paraspeckle associated eler	nents
Paraspeckles as rhythmic nuclear mRNA anchorages responsible for circadian	101 gene
expression	101
Introduction	102
Résultats	103
Conclusion	104
Recherche de déterminants moléculaires impliqués dans la rétention d'ARNm	par les
paraspeckles	
Introduction	
Matériel et méthodes	136
Résultats :	139
Discussion	145
Discussion et Perspectives	148
1 - Neat1 RNA pull-down : mise au point et apport de la technique	149
Apport du RNA pull-down pour l'étude des IncRNAs	149
Mise au point du protocole de RNA pull-down:	150
2 - Structure des paraspeckles dans les cellules hypophysaires	153
Caractérisation des paraspeckles dans la lignée GH4C1	153
Structure des paraspeckles dans la lignée GH4C1	154
3 - Neat1 et la rétention nucléaire des ARNs	155
Rétention d'ARNs par les paraspeckles dans la lignée GH4C1 :	155
Etude des motifs permettant la rétention des ARNs par les paraspeckles	155
4 – Les paraspeckles dans l'oscillateur hypophysaire	158
Rythmicité circadienne d'expression des éléments constituants des paraspeckles	158
Conséquences de la rythmicité circadienne de formation des paraspeckles	159
5 – Neat1, organogénèse et rythmicité circadienne :	162
Conclusion générale	164
Bibliographie	166
	-

## **Introduction générale**

Du fait de la rotation de la terre, les organismes vivants évoluent dans un environnement fluctuant quotidiennement. La capacité d'anticiper et de se synchroniser sur ces changements environnementaux représente un avantage évolutif majeur qui a conduit les espèces vivantes à développer un système d'horloge interne endogène oscillant sur une période d'environ 24h. Ce système, dit circadien, qui permet de générer des rythmes, est coordonné par une horloge centrale localisée chez les mammifères dans le noyau suprachiasmatique (NSC) de l'hypothalamus. Le NSC intègre les signaux en provenance de l'environnement permettant la mise à l'heure du système sur les informations environnementales. On sait que la rythmicité circadienne repose sur un mécanisme intracellulaire autonome, que l'on définit comme « l'oscillateur moléculaire » de la cellule. Dans chaque cellule, l'expression périodique de gènes spécifiques codant pour des protéines impliquées dans des boucles d'autorégulation transcriptionnelle négative permet de générer des oscillations circadiennes (Vitaterna et al 1994; Shearman et al 2000). Ce mécanisme n'est pas l'apanage des cellules de l'horloge centrale mais existe également dans toutes les cellules de l'organisme. La synchronisation des cellules au sein d'un tissu est nécessaire à l'émergence d'une rythmicité circadienne à l'échelle des organes. Cette synchronisation est assurée par l'horloge centrale au travers de signaux neuronaux, endocriniens et comportementaux (Buijs, & Kalsbeek 2001; Yoo et al 2004; Guo et al 2006).

La rythmicité circadienne concerne une multitude de processus depuis le métabolisme cellulaire jusqu'aux activités physiologiques et comportementales. Chez les mammifères, la rythmicité de ces processus est soutenue par l'expression rythmique de nombreux gènes. L'évolution des techniques d'étude du génome a permis de déterminer et de revoir à la hausse l'importance quantitative de l'ensemble des gènes rythmiques, c'est-à-dire le transcriptome rythmique de l'organisme. On estime aujourd'hui que selon le type cellulaire, le tissu ou les conditions environnementales, entre 8% à 40% des gènes sont exprimés de manière rythmique dans un organe. On admettait jusqu'à récemment que la plupart des gènes rythmiques présentaient une

transcription rythmique, cependant des études récentes ont révélé que la rythmicité de nombreux ARNs rythmiques est contrôlée au niveau post-transcriptionnel (Koike et al 2012; Menet et al 2012; Le Martelot et al 2012; Lück et al 2014; Atger et al 2015) . Des données de la littérature montrent que chacune des étapes qui contrôlent le devenir des ARNs à l'intérieur de la cellule peut être soumise à un rythme circadien (Harms et al 2004; Green 2017) (torres et al. 2018; sous presse). Rechercher quels sont ces mécanismes largement méconnus qui s'exercent après la transcription et qui sont impliqués dans la régulation des rythmes circadiens est un des enjeux actuels en chronobiologie.

Au cours de cette étude nous avons été amenés à nous intéresser à l'un de ces mécanismes dans une lignée de cellules hypophysaires, la lignée GH4C1. Cette lignée de type somatolactotrope exprime l'hormone de croissance (Gh) et la prolactine (Prl). Des travaux antérieurs de l'équipe avaient permis de caractériser dans cette lignée cellulaire la présence d'un oscillateur circadien autonome et fonctionnel (Becquet et al 2014). En recherchant les mécanismes moléculaires permettant l'expression rythmique du gène de la Prl, l'équipe avait montré que dans cette lignée cellulaire deux protéines, les protéines NONO et SFPQ impliquées dans l'expression rythmique de la Prl (Guillaumond et al 2011) suivent un rythme circadien d'expression. Ces protéines sont multifonctionnelles et entrent dans la composition de structures nucléaires appelées les paraspeckles. Les paraspeckles, qui correspondent au sous-domaine nucléaire le plus récemment identifié, se forment autour d'un long ARN non codant, Neat1, auquel s'associent non seulement NONO et SFPQ mais aussi de nombreuses autres protéines (Fox et al 2002; Clemson et al 2009; Sasaki et al 2009; Sunwoo et al 2009) . Différentes études s'accordent à montrer que les paraspeckles sont capables de retenir dans le noyau certains ARNs participant ainsi au contrôle de leur export cytoplasmique (Prasanth et al 2005; Chen et al 2008; Mao et al 2011) . La rétention nucléaire de ces ARNs impliquerait la reconnaissance de motifs particuliers présents dans la région 3'UTR de ces ARNs qui correspondent à deux séquences Alu inversées (IRAlus) qui vont s'apparier pour former une boucle d'ARN double brin.

Deux des protéines appartenant aux paraspeckles, les protéines NONO et SFPQ, étant exprimées de manière rythmique, nous avons émis l'hypothèse selon laquelle la formation des paraspeckles à l'intérieur du noyau des cellules pouvait suivre une rythmicité circadienne. Dans cette hypothèse, ces sous-domaines nucléaires pourraient induire la rythmicité de la rétention nucléaire des ARNs qu'ils reconnaissent.

Avec pour but de rechercher l'implication des paraspeckles dans l'expression rythmique d'ARNs, nous avons en premier lieu développé une approche de « RNA pull down » permettant d'identifier les ARNm associés à un long ARN non codant (torres et al. 2018, JoVE, sous presse). Cette approche appliquée à Neat1 permet d'isoler et d'identifier les ARNs potentiellement retenus par les paraspeckles.

Nous avons dans un deuxième temps montré que les autres composants des paraspeckles dont Neat1 ainsi que le nombre des paraspeckles présentent une rythmicité circadienne d'expression et d'association et que grâce à cette expression circadienne, ils peuvent contrôler la rythmicité des ARNm qui leur sont associés (Torres et al 2016; Torres et al 2017).

L'ensemble de notre étude a permis de décrire un nouveau mécanisme posttranscriptionnel original impliqué dans le fonctionnement du système circadien et d'attribuer un rôle dans la physiologie de l'hypophyse, à un long ARN non-codant qui est l'élément structural de sous-domaines nucléaires. **Revue de littérature** 

# I – Les rythmes circadiens

### **1. Les rythmes biologiques**

### 1.1. Importance des rythmes biologiques

Du fait de la rotation de la derre sur son axe et de sa course autour du soleil, l'environnement dans lequel les organismes vivent sur terre est en perpétuel changement. Cependant ces modifications de l'environnement présentent l'avantage d'être régulières et peuvent donc être anticipées. Le passage du jour à la nuit est un cycle régulier duquel va dépendre les phases d'activité et de repos de nombreuses espèces. Durant les périodes d'activité, les animaux investissent plus d'énergie dans leurs stratégies de survie et de reproduction alors que les phases de repos sont plutôt consacrées à des processus de régénération de l'organisme. En plus de ces alternances d'activité journalières, les animaux doivent fonctionner en tenant compte des changements de saisons qui vont avoir pour conséquences majeures des variations drastiques du climat et une modification de la quantité de ressources disponibles dans l'environnement. Pour survivre à des périodes hostiles, les animaux doivent posséder des réserves d'énergie leur permettant de tenir durant ces saisons et calquer leur période de reproduction sur les saisons les plus favorables à la survie de leur descendance. Les espèces capables de synchroniser leur comportement et leur métabolisme sur les périodes les plus favorables à leur succès auraient alors une meilleure capacité à survivre que des espèces incapables de s'adapter aux modifications de leur environnement. Pour pouvoir anticiper ces changements qui dépendent de l'alternance jour/nuit et des changements de saisons, les organismes vivants doivent donc être capables de mesurer de manière précise le passage du temps et d'adapter leur comportement en fonction des heures du jour, ou des saisons de l'année. Disposer d'une telle capacité présente un avantage évolutif certain au point que la majorité des espèces vivantes aujourd'hui, des cyanobactéries jusqu'à l'être Humain, possèdent des horloges biologiques endogènes leur permettant cette mesure du temps qui passe (Harmer et al 2001; Bell-Pedersen et al 2005). Ces horloges vont impacter toutes les fonctions d'un organisme, aussi bien le métabolisme cellulaire, les fonctions physiologiques que le comportement. Chez l'Humain par exemple, la

libération de glucocorticoïdes augmente en fin de nuit avec un pic le matin ce qui permet de mobiliser des ressources énergétiques en anticipation de la phase d'éveil alors que la sécrétion d'hormone de croissance atteindra plutôt son pic au cours de la nuit pour une efficacité maximale de réparation et de renforcement des os durant la phase de repos (Takahashi et al 1968).

Si le bon positionnement temporel des différents processus métaboliques, physiologiques et comportementaux représente un moyen d'optimiser l'ensemble des fonctions vitales, à l'inverse un mauvais fonctionnement de cette horloge interne va perturber cet équilibre. Il est aujourd'hui reconnu que les perturbations des rythmes qu'elles soient d'origine génétique ou environnementale (par exemple le *jet lag* ou les horaires décalés) sont des facteurs à risque dans le développement de pathologies ou des facteurs aggravants dans le cas de pathologies préexistantes.

Chez l'humain il a été montré que les perturbations environnementales des rythmes causent une augmentation des risques de développement de cancers (Wegrzyn et al 2017), de syndromes métaboliques (obésité, diabètes) (Mayeuf-Louchart et al 2017) et de maladies cardiovasculaires (Karlsson et al 2001). Les perturbations de l'horloge sont aussi fortement associées aux maladies neurodégénératives comme Alzheimer, Parkinson ou la maladie d'Huntington qui ont en commun des perturbation de la rythmicité de fonctions comportementales (modification des cycles veille/sommeil), des fonctions physiologiques et cellulaires (Hood, & Amir 2017). Les dysfonctionnements de l'horloge interne apparaissent tôt dans la genèse de ces pathologies et deviennent alors des facteurs aggravants qui participent à la dégradation de l'état général du malade.

### 1.2. Les rythmes circadiens

### **1.2.1.** Découverte des rythmes circadiens

La rythmicité des fonctions biologiques a été découverte longtemps avant l'introduction du concept d'horloges biologiques endogènes. L'hibernation de certaines espèces, la saisonnalité des périodes de reproduction chez les animaux ainsi que de la floraison et de la fructification des plantes sont autant de phénomènes observés par l'homme depuis des siècles. Chez l'homme, d'anciens traités de médecine attestent de la connaissance de variations physiologiques qui se produisent au cours d'une journée. Ainsi John Wren, publie dans son traité d'herbologie en 1632, que le sang présente des variations d'humeur au cours de la journée.

« Blood reyneth 6 h from 9 o'clock in ye night till 3 in the morne, cholic from 3 in the morne till 9: melancholy from 9 till 3 in the even: flegme from 3 till 9 o'clock in the night ».

Cependant la nature endogène de certains rythmes ne sera rapportée qu'au XVIIIème siècle grâce à des études faites sur les plantes. En 1729, le mathématicien, astronome et géophysicien français Jean-Jacques d'Ortous de Mairan observe que les feuilles de la sensitive (*Mimosa pudica*) ont un mouvement calqué sur l'alternance des jours et des nuits. En plaçant ses plants dans des caves obscures il constate que les mouvements des feuilles continuent même en l'absence d'indications temporelles extérieures et en déduit que la rythmicité du mouvement des feuilles est indépendante des signaux environnementaux. Il faudra ensuite attendre 1939 pour que M. Johnson démontre que chez la souris aussi, la rythmicité de l'activité locomotrice est rythmique et persistante en l'absence de signaux environnementaux.

### 1.2.2. Définition des rythmes circadiens

L'observation chez les plantes ou les mammifères de rythmes persistants, même en l'absence de signaux temporels en provenance de l'environnement, a fait émerger la notion d'endogénie des rythmes. Cette propriété peut être testée par des expériences d'isolement du modèle étudié dans un contexte d'obscurité constante par exemple. Dans ces conditions, il est aussi possible de mesurer la période du rythme étudié, on parle alors de période de « free-running » d'un rythme. Pour un même paramètre étudié, la période peut varier entre les espèces mais aussi entre les individus. Il existe une grande diversité de rythmes endogènes observés en biologie. Trois grandes classes de rythmes endogènes sont établies sur le critère de leur période. La catégorie de rythmes la plus étudiée est celle des rythmes endogènes journaliers d'une période variant entre 20h et 28h qui ont été appelés rythmes circadiens (du latin :



### Figure 1 : Modélisation mathématique d'un rythme biologique

Un rythme biologique peut être analysé par la méthode du cosinor. Cette méthode modélise les variations d'une variable biologique par une fonction sinusoïdale dont l'équation est :

Y (t) = M + A Cos ( $\omega$  t +  $\Phi$ ) (t est le temps;  $\omega$  est la fréquence angulaire = 2  $\pi$  /  $\tau$ ) Cette équation définit quatre paramètres: La période  $\tau$ , représente la durée d'un cycle complet de la variable biologique. L'acrophase  $\Phi$  et la bathyphase, représentent respectivement les moments où la variable biologique est à son amplitude maximale ou minimale. Le mésor M, représente la valeur moyenne de la variable biologique rapportée à la durée de la période. L'amplitude A, représente la différence des valeurs entre le pic et le Mésor. *circa*=proche et *dies*=jour) par Franz Halberg en 1959. Il existe cependant aussi des rythmes infradiens d'une période supérieure à 28h et des rythmes ultradiens aux périodes inférieures à 20h.

Lorsque les rythmes circadiens sont synchronisés par les signaux environnementaux (ou Zeitgebers) comme l'alternance jour/nuit, ils présentent une période de 24h, on parle alors de rythmes nycthéméraux. Le processus de synchronisation d'une période endogène sur un rythme externe est désigné comme l'entrainement de l'oscillateur et permet de maintenir en phase les rythmes internes avec les rythmes externes. Chez la plupart des espèces, la photopériode constitue le signal majeur d'entrainement des rythmes circadiens pour les synchroniser avec les cycles jour/nuit de l'environnement. En plus de leur période les rythmes circadiens se définissent aussi par leur phase. La phase désigne le positionnement d'un rythme endogène par rapport à un autre rythme endogène ou par rapport au cycle jour/nuit de l'environnement. Par exemple les espèces diurnes ont une activité motrice en phase avec le jour, alors que les espèces nocturnes vont avoir une activité motrice en phase avec la nuit.

Enfin, les rythmes circadiens se définissent aussi par leur amplitude. L'amplitude représente le degré de variation du paramètre mesuré par rapport à son niveau moyen ou mésor (Fig1).

### 2. Organisation du système circadien

### 2.1. L'horloge centrale chez les mammifères

### 2.1.1. Mise en évidence de l'horloge centrale dans le noyau suprachiasmatique

L'étude de V.Critchlow, en 1963 montre que chez le rat, une lésion de la région suprachiasmatique de l'hypothalamus induit une désynchronisation de la phase de l'oestrus chez les femelles (Critchlow 1963). Ces résultats indiquant un rôle de la région suprachiasmatique de l'hypothalamus dans le fonctionnement des rythmes circadiens sont encouragés par la découverte chez le rat d'une projection directe rétino-hypothalamique dans le Noyau Suprachiasmatique (NSC) de l'hypothalamus (Moore, & Lenn 1972). La même année sont publiées deux études de lésion de NSC chez le rat qui montrent d'une part que le NSC est impliqué dans la génération du rythme circadien de sécrétion de la corticostérone (Moore, & Eichler 1972), et d'autre part que le NSC est nécessaire au maintien d'une activité circadienne dans la prise de boisson et le comportement locomoteur (Stephan, & Zucker 1972). Chez le hamster doré, la lésion du NSC induit une activité motrice sporadique en conditions d'alternance jour/nuit (L/D) et arythmique en obscurité permanente (D/D). Ces résultats indiquent l'implication du NSC dans les fonctions rythmiques endogènes.

Par la suite, des équipes ont essayé de restaurer les processus circadiens dépendants du NSC par des greffes d'hypothalamus antérieur fœtal. Les greffons ont été transplantés dans le 3éme ventricule de rats ayant subi une lésion des NSC. Après la greffe une récupération des comportements circadiens moteurs perdus suite à la lésion du NSC est observée (Drucker-Colín et al 1984; Sawaki et al 1984) . Chez le hamster, cette récupération de rythmes circadiens s'observe aussi en condition d'obscurité permanente démontrant le rétablissement d'une rythmicité endogène indépendante de la photopériode. Le bon fonctionnement de la greffe est corrélé avec l'expression du polypeptide intestinal vasopressine (VIP), du neuropeptide Y (NPY) et de la vasopressine (VP) qui sont des neuropeptides exprimés dans le NSC de hamster non lésés (Lehman et al 1987). Enfin ces expériences de greffes montrent aussi l'importance des projections du NSC dans le maintien des rythmes. En effet il a été



В



### Figure 2 : Localisation anatomique et architecture du noyau suprachiasmatique (NSC)

A: Localisation anatomique du NSC dans l'hypothalamus

B: A gauche, l'organisation du NSC en deux subdivisions est schématisée. A droite, les populations neuronales majeures (gros caractères) et accessoires (petits caractères) présentes dans les deux subdivisions du noyau chez le rat sont mentionnées. 3V : troisième ventricule

constaté que les greffes ayant les meilleurs résultats étaient celles formant le plus de contact avec des projections en provenance du tractus retino-hypothalamique et développant le plus de projections vers les structures environnantes. Néanmoins ce modèle de greffe ne permet pas de récupérer intégralement les fonctions dépendantes du NSC, et les rythmes hormonaux ne sont pas restaurés suite à une greffe d'hypothalamus antérieur fœtal contenant les NSC (Lehman et al 1987).

### 2.1.2. Anatomie et organisation du noyau suprachiasmatique

Le noyau suprachiasmatique (NSC) est une structure composée de deux noyaux bilatéraux situés au dessus du chiasma optique. Le NSC se compose d'environ 20 000 neurones ce qui fait de lui un petit noyau en comparaison des autres noyaux de l'hypothalamus. Il se compose de petits neurones qui sont majoritairement GABAergiques et de cellules gliales, en particulier des astrocytes. Le NSC est une structure hétérogène. Il peut être divisé en deux compartiments, une région ventrolatérale et une région dorso-médiane, dont les neurones se distinguent par leur forme, les peptides qu'ils expriment et les fonctions qui leur sont attribuées dans le fonctionnement du système circadien (Fig2).

Les neurones de la partie ventro-latérale sont majoritairement petits et de forme ovale ou ronde et leur nombre représente approximativement un quart des neurones du NSC. Chez le rat, la majorité des neurones de cette région expriment le peptide vasoactif intestinal (VIP) et le peptide gastro-intestinal (GRP) (van den Pol, & Tsujimoto 1985). Ces neurones présentent des projections synaptiques vers leurs voisins dans la région ventro-latérale du NSC mais aussi vers les neurones de la partie dorso-médiane du NSC.

La région dorso-médiane du NSC se compose de petits neurones de forme ronde, présents en plus grande densité que ceux de la partie ventrale. Le peptide principalement exprimé par ces neurones est le peptide arginine vasopressine (AVP).

Tous les neurones du NSC présentent une activité électrique circadienne (Green, & Gillette 1982). Cependant, dissociés et mis en culture les neurones du NSC ne sont pas en phase entre eux (Welsh et al 1995). Les neurones du NSC sont donc autant d'oscillateurs autonomes dans leur fonction circadienne. Cependant, individuellement

la rythmicité de ces oscillateurs est asynchrone et leur capacité à rythmer est plus faible. La mesure de la rythmicité de l'activité neuronale dans des neurones de NSC de souris montre que plus les neurones sont isolés dans des cultures de neurones dissociés de faible densité, et plus leur oscillation circadienne devient instable (Webb et al 2009; Herzog et al 2004). En revanche les neurones du NSC lorsque ils sont couplés dans leur réseau tissulaire, expriment des rythmes plus précis, plus stables, et de plus grande amplitude. L'association en réseau des neurones du NSC apporte une plus grande robustesse à l'oscillateur présent dans le NSC (Abraham et al 2010; Buhr et al 2010; Nakamura et al 2002; Liu et al 2007; Evans et al 2012; Ko et al 2010) .

#### 2.1.3. Fonctionnement moléculaire de l'oscillateur circadien

Dans les années 60, des travaux précurseurs montrent des perturbations des rythmes circadiens après utilisation d'agents perturbateurs de la synthèse des ARNs (Karakashian, & Hastings 1962), ou inhibiteurs de la synthèse des protéines (Feldman 1967). Ces travaux conduits chez des formes primaires d'eucaryotes indiquent l'existence de mécanismes moléculaires impliqués dans la genèse des rythmes circadiens.

Une étape clé pour comprendre ces mécanismes a été la génération de mutations chez la drosophile afin de passer au crible les gènes impliqués dans les rythmes circadiens. De cette étude ressort trois mutants en particulier dont le rythme de 24h est dramatiquement endommagé ; l'un est arythmique, le second a une période de rythmicité inférieure à 24h et le dernier une période plus longue. Les trois mutations impliquent un gène présent dans le chromosome X appelé plus tard le gène Période (Per) (Konopka, & Benzer 1971). Par homologie, l'équivalent de Per chez l'humain et la souris est retrouvé (Tei et al 1997; Sun et al 1997) . Il existe trois homologues de Per chez la souris et l'humain (Per1, Per2, Per3) qui possèdent, comme chez la drosophile, un domaine Per-Arnt-Sim (PAS) et ont une expression circadienne dans le NSC. Chez la souris, la mutation du domaine PAS des gènes Per induit un raccourcissement de la période circadienne en alternance jour/nuit, et une abolition de la rythmicité en obscurité continue (Zheng et al 1999; Cermakian et al 2001).



### Figure 3 : Mécanisme moléculaire des oscillations circadiennes chez les mammifères

Le mécanisme moléculaire à la base de la rythmicité circadienne repose sur des boucles d'autorégulation de la transcription impliquant le complexe dimérique activateur constitué des protéines CLOCK/BMAL1. Ce complexe active la transcription des gènes Per, Cry, Rev-erb et ROR via des séquences E-Box présentes dans leur promoteur. Les protéines PER et CRY associées en hétérodimères vont rétroinhiber leur propre transcription, formant une première boucle de rétrocontrôle transcriptionnel négatif. Les récepteurs nucléaires REV-ERB $\alpha$ , $\beta$  et ROR $\alpha$ , $\beta$ , $\gamma$  respectivement inhibent ou activent la transcription de Bmal1 en se liant sur des séquences ROREs, formant une seconde boucle d'autorégulation. Au niveau post-transcriptionnel, les caséines kinases (CK1) participent au maintien de la période circadienne des boucles d'autorégulation transcriptionnelle. Ces boucles ont une période autour de 24h.

Chez les mammifères, une étude menée chez la souris permet dans les années 90 de découvrir le gène Clock. La perte de fonction de Clock par mutation allonge la période circadienne et abolit la persistance des rythmes en conditions constantes (Vitaterna et al 1994). Clock code pour un facteur de transcription de la famille des facteurs de type basic helix-loop-helix (bHLH) à domaine PAS (King et al 1997). La protéine CLOCK présente des sites de liaison à l'ADN et peut s'associer en dimères à d'autres protéines. La recherche de partenaires protéiques a montré que CLOCK s'associe à BMAL1 (ou MOP3), une autre protéine de la famille des facteurs bHLH-PAS. Chez la souris, la perte d'expression de BMAL1 induit une perturbation de l'activité locomotrice circadienne en condition d'alternance jour/nuit (LD) ainsi qu'une abolition complète et immédiate de la rythmicité en obscurité constante (DD) (Bunger et al 2000). BMAL1 est donc un gène majeur dans le fonctionnement de l'oscillateur circadien.

L'hétérodimère BMAL1:CLOCK a la capacité de se fixer sur des séquences de type Ebox situées en cis- dans le promoteur de certains gènes et induit ainsi la transcription de ces gènes (Gekakis et al 1998). Les gènes Period (Per1, Per2 et Per3) et Cryptochrome (Cry1 et Cry2) sont régulés par l'hétérodimère BMAL1:CLOCK. Les protéines correspondantes sont exprimées dans le cytoplasme en antiphase avec l'expression de BMAL1. CRY1 et CRY2 s'associent ensuite avec les protéines PER et permettent ainsi leur translocation vers le noyau (Kume et al 1999). Dans ce complexe protéique, il a été montré que les protéines CRY sont nécessaires à la stabilisation de PER (Lee et al 2001). Chez la souris, l'absence de CRY1 et CRY2 se traduit par une arythmie du comportement (Van Der Horst et al 1999) à laquelle est corrélée une perte de la rythmicité d'expression des ARNm de Per1 et Per2 (Okamura et al 1999; Vitaterna et al 1999) . En effet le complexe PER/CRY, par une interaction protéine/protéine, inhibe dans le noyau le complexe BMAL1:CLOCK (Shearman et al 2000; Sato et al 2006) réduisant ainsi leur propre transcription et formant une première boucle de rétro-inhibition dont la période est d'environ 24h (Fig3).

Une autre cible du complexe BMAL1:CLOCK est le gène du récepteur orphelin Rev-erba dont le promoteur contient plusieurs séquences E-BOX (Triqueneaux et al 2004). Du fait de la première boucle de rétro-inhibition décrite précédemment, le complexe BMAL1:CLOCK va induire la transcription de Rev-erba avec une rythmicité circadienne. Dans le noyau, la protéine REV-ERBa s'associe de manière transitoire à l'hème, de la famille des porphyrines contenant du fer qui servent de support aux enzymes du métabolisme oxydatif et aux facteurs de transcription. Sous cette forme, REV-ERB $\alpha$  peut se lier à une séquence consensus de type RORE qui se trouve être exprimée dans le promoteur du gène codant pour Bmal1. La fixation de REV-ERB $\alpha$  au promoteur de Bmal1 induit alors l'inhibition de la transcription de Bmal1 et par conséquent une réduction de l'expression de la protéine BMAL1 (Preitner et al 2002; Yin et al 2007). Chez des souris déficientes en REV-ERB $\alpha$ , il a été montré la conservation d'un comportement rythmique avec cependant une réduction de la période des rythmes comportementaux ainsi qu'une plus grande irrégularité dans la longueur des périodes (Preitner et al 2002). REV-ERB $\alpha$  est donc un élément indispensable au contrôle de la durée et de la régularité de la première boucle de rétrocontrôle.

Suite à des études génomiques des gènes à expression rythmique dans le NSC, le cœur et le foie de souris (Panda et al 2002; Akhtar et al 2002; Storch et al 2002; Ueda et al 2002), un autre régulateur de Bmal1 a pu être découvert. Les protéines ROR, de la famille des récepteurs orphelins de type acides rétinoïques, présentent comme REV-ERB $\alpha$  un domaine de liaison à la séquence consensus RORE. La fixation de ROR ( $\alpha$ ,  $\beta$  et  $\gamma$ ) à la séquence RORE dans le promoteur de Bmal1 induit une augmentation de la transcription de Bmal1 abolie lorsque la séquence RORE est mutée (Guillaumond et al 2005; Sato et al 2004). Une déficience en ROR $\alpha$  induit « in vivo » chez la souris un comportement moteur aberrant ainsi qu'une réduction de la stabilité de la rythmicité circadienne comportementale ; « in vitro » en culture cellulaire, la déficience en ROR $\alpha$  réduit l'amplitude des rythmes d'expression de Bmal1 (Akashi, & Takumi 2005).

Les protéines REV-ERB et ROR, de par leur affinité pour les séquences ROREs, sont donc en compétition pour la fixation sur le promoteur de Bmal1 (Guillaumond et al 2005). Cette balance entre les effets de REV-ERB et ROR participe à la régulation de la rythmicité d'expression de Bmal1. Ce système forme une seconde boucle de rétrocontrôle dont le principal rôle est de stabiliser la période de l'oscillateur circadien moléculaire (Fig3).

A ces deux boucles de rétrocontrôle s'ajoutent des mécanismes post-traductionnels qui renforcent et maintiennent l'oscillateur à une période de 24h.

Les caséines kinases, de la famille des protéines kinases, jouent un rôle important dans la stabilité de ces boucles. Chez le hamster Syrien, il a été reporté qu'en réduisant par mutation l'efficacité de l'enzyme CK1 $\varepsilon$  on induit une modification de la période des rythmes circadiens (Lowrey et al 2000). Chez la souris il a été montré que les protéines PER1 et PER3 sont des substrats des caséines kinases CK1 $\varepsilon$  et CK1 $\delta$  (Lee et al 2001; Akashi et al 2002). La phosphorylation par ces kinases de PER1 et PER3 régule leur vitesse de dégradation, permettant ainsi d'assurer que leur cycle de renouvellement soit cohérent avec la périodicité circadienne de l'oscillateur moléculaire (Eide et al 2005). De plus CK1 $\varepsilon$  et CK1 $\delta$  sont aussi capables d'induire la translocation au noyau de PER3 en modelant ses signaux de localisation nucléaire. Ces résultats sont en accord avec l'observation chez l'humain d'une mutation du site de liaison de CK1 $\varepsilon$  à la protéine hPER2 causant une hypo-phosphorylation de la protéine. Cette mutation est trouvée dans des cas de formes familiales de syndrome d'avance de phase de sommeil (Toh et al 2001).

En plus de cet effet sur la phosphorylation des protéines PER, il a aussi été montré qu'à travers leur liaison à PER, les protéines CRY1 et CRY2 sont aussi phosphorylées par CK1ε (Eide et al 2002) et que BMAL1 est aussi une cible de CK1ε qui a un effet positif sur l'activation des promoteurs dépendants de BMAL1 (Eide et al 2002).

Chez les mammifères, des mutations de ces deux caséines kinases se traduisent par une réduction de la période du cycle circadien (Ralph, & Menaker 1988; Lowrey et al 2000; Xu et al 2005), illustrant ainsi le rôle prépondérant des caséines kinases CK1ε et CK1δ dans le maintien de la période de l'oscillateur circadien (fig3).

Le fonctionnement de cet oscillateur moléculaire circadien va fortement influencer l'activité électrique rythmique des neurones du NSC. L'enregistrement par multiélectrodes de neurones du NSC dissociés montre qu'une mutation du gène Clock impacte la décharge électrique des neurones du NSC, la forme hétérozygote de la mutation induisant un allongement de la période du rythme de décharge, et la forme homozygote induisant une arythmie de cette décharge neuronale (Herzog et al 1998). De la même manière, une mutation de la CK1ɛ raccourcit le rythme de l'activité électrique spontanée des neurones du NSC et est associée à un raccourcissement de la période de l'activité locomotrice (Liu et al 1997).

Ce modèle d'oscillateur générateur de rythmes circadiens est fortement conservé entre les espèces. En effet, si les gènes impliqués sont peu conservés du point de vue évolutif, le mécanisme de boucles d'autorégulation de transcription et de traduction est lui retrouvé dans la grande majorité des organismes rythmiques (Glossop et al 1999; Lee et al 2000)( pour review voir (Mackey 2007)).

### 2.2. Synchronisation externe du système circadien

### 2.2.1 Voies anatomiques d'entrée sur le NSC

L'entrainement de l'oscillateur du NSC par des inputs extérieurs se fait majoritairement par des signaux neuronaux directs ou indirects. Ces signaux sont transmis principalement par trois voies de synchronisation. Deux voies transmettent des signaux de synchronisation dits photiques depuis la rétine ; la première qui forme le tractus rétino-hypothalamique est une voie directe, et la seconde, est une voie indirecte qui fait relais dans le thalamus pour se terminer par le tractus géniculohypothalamique (Fig4). La troisième voie véhicule les signaux de synchronisation dits non-photiques et provient principalement du système sérotoninergique (Fig4).

La projection du tractus rétino-hypothalamique (TRH) sur le NSC a été montrée chez le rat par transport antérograde et autoradiographie. En injectant des protéines marquées dans la partie postérieure de l'œil il a été observé un transport de ces protéines depuis l'œil vers la partie ventrale du NSC de l'hypothalamus. Suite à une énucléation des yeux, une dégénération des terminaisons axonales présentes dans le NSC est constatée, confirmant ainsi l'existence de fibres afférentes directes de l'œil vers le NSC (Moore, & Lenn 1972). La même étude réalisée chez différentes espèces de mammifères (opossum, hérisson, musaraigne, chat, ouistiti, macaque) confirme que la projection directe des fibres depuis la rétine vers le NSC est une caractéristique commune à la classe des mammifères (Moore 1973). Les axones à glutamate et à pituitary adenylate cyclase-activating peptide (PACAP) du TRH forment de nombreuses synapses avec les dendrites des neurones VIP et GRP du NSC aussi bien chez le rat que chez le hamster (Tanaka et al 1997; Aioun et al 1998; Ibata et al 1989; Abrahamson, & Moore 2001) .



### Figure 4: Voies afférentes du noyau suprachiasmatique

L'oscillateur présent dans le noyau suprachiasmatique est entrainé par des signaux photiques et non photiques.

Les voies photiques (en violet) véhiculent des signaux en provenance des cellules ganglionnaires de la rétine soit directement par le tractus rétino-hypothalamique (TRH), soit indirectement par le tractus géniculo-hypothalamique (TGH) dont le relais se fait dans le feuillet intergéniculé latéral du thalamus.

Les signaux non-photiques sont principalement véhiculés par la voie sérotoninergique (5-HT) (en vert) par des neurones dont les corps cellulaires sont localisés dans les noyaux dorsal et médian du Raphé.

Ces trois voies projettent principalement sur les neurones de la région ventro-latérale du noyau suprachiasmatique.

La seconde voie d'afférences photiques vers le NSC est le tractus géniculohypothalamique (TGH). Depuis la rétine, des projections transmettent l'information photique au feuillet intergéniculé (IGL) situé dans le corps géniculé latéral du thalamus. L'IGL innerve à son tour le NSC au niveau de la partie ventro-latérale (Hisano et al 1988), formant ainsi le TGH. Chez le rat, les fibres du TGH expriment le neuropeptide Y (NPY) et l'acide  $\gamma$ -aminobutyrique (GABA). Les neurones NPY/GABA de l'IGL forment des contacts synaptiques principalement avec les neurones VIP de la partie ventro-latérale du NSC (Francois-Bellan et al 1990).

La troisième voie majeure de projection sur le NSC est la voie sérotoninergique, qui est non-photique. Chez le rat, les informations sérotoninergiques proviennent des noyaux dorsal et médian du Raphé et projettent sur la région ventro-latérale du NSC. Les fibres sérotoninergiques établissent des synapses avec les neurones à VIP de la partie ventrolatérale du NSC (Kiss et al 1984; Bosler, & Beaudet 1985). Cependant, une partie des fibres sérotoninergiques ne forme pas de synapses dans le NSC, suggérant aussi une action paracrine de la sérotonine dans le NSC (Bosler 1989).

### 2.2.2. Synchronisation photique

La synchronisation du NSC par l'information lumineuse est initiée au niveau de la rétine. Les cellules ganglionnaires de la rétine projettent vers le NSC. Des études indiquent que ces cellules contiennent deux photopigments, la mélanopsine et les cryptochromes, qui sont impliqués dans la transmission de l'information lumineuse vers l'horloge (Cermakian, & Sassone-Corsi 2002).

L'information codée par les cellules ganglionnaires est ensuite transportée vers le NSC par le TRH et le signal est transmis par la libération de glutamate et de PACAP au niveau des terminaisons axonales du TRH (Vindlacheruvu et al 1992). Le TRH va ainsi principalement donner l'information lumineuse aux neurones de la partie ventro-latérale du NSC (Abrahamson, & Moore 2001; Lokshin et al 2015) . En effet, l'enregistrement de l'activité extracellulaire des neurones du NSC aussi bien chez le rat que chez le hamster a permis de montrer que suite à des stimulations lumineuses de la

rétine, ce sont les neurones de la partie ventro-latérale du NSC qui répondent par une modification de leur activité électrique (Meijer et al 1986).

La libération de glutamate et de PACAP va impacter les neurones du NSC en activant plusieurs voies de signalisation. Principalement, le glutamate va activer les voies du monoxyde d'azote (NO) et des Kinases dépendantes du Calcium/Calmodulin (CaMK), alors que PACAP va activer la voie des MAPK (mitogen-activated protein kinase). Ces voies de signalisation activées différemment selon la phase de la photopériode, vont notamment permettre la phosphorylation du facteur de transcription CREB (cAMP responsive element binding protein). CREB est spécifiquement phosphorylé pendant la nuit et va alors réguler la transcription de nombreux gènes dans le NSC dont Per et c-Fos (Ginty et al 1993; Ding et al 1997; Schurov et al 1999; Gau et al 2002) .

Les gènes Per et c-Fos sont parmi les principaux gènes qui répondent à l'information lumineuse dans la partie ventro-latérale du NSC (Gau et al 2002). Le degré d'induction de Per1 est fortement corrélé à la fréquence de décharge neuronale (Yan, & Okamura 2002; Kuhlman et al 2003; Yan, & Silver 2004) et l'inhibition de Per1 induit la perte des modifications de la phase des rythmes comportementaux par la lumière (Akiyama et al 1999). Plus tard, l'utilisation de souris Per1::GFP a permis de montrer que les neurones GRP n'expriment pas Per1 de manière circadienne contrairement à d'autres populations de neurones du NSC (Karatsoreos et al 2004). Dans ces neurones, per1 est entrainé par la photopériode ce qui synchronise ensuite le reste de l'oscillateur moléculaire.

La partie ventro-latérale est donc la partie du NSC majoritairement impliquée dans l'intégration des signaux lumineux en provenance de la rétine. Les connections synaptiques entre les neurones de la partie ventro-latérale et ceux de la partie dorsomédiane vont permettre à cette information d'être propagée à l'ensemble du NSC.

#### 2.2.3. Synchronisation non-photique

En dehors des signaux de synchronisation photiques, des signaux non-photiques participent aussi à l'ajustement de la phase des neurones du NSC sur les informations externes.

En particulier, la voie sérotoninergique participe à la synchronisation de la rythmicité circadienne dans le NSC. La rupture de la voie sérotoninergique, par coupure des NSC utilisation d'inhibiteurs projections sérotoninergiques du ou par pharmacologiques de la sérotonine (5-HT), induit une altération des rythmes circadiens comportementaux et endocriniens (Rea 1998). L'activation de certains types de récepteurs de la 5-HT dans le NSC a une influence sur la phase des rythmes d'activité neuronale ou motrice seulement lorsqu'ils sont activés durant le jour (Edgar et al 1993; Prosser et al 1993). En revanche leur activation durant la nuit n'entraine pas de modification de la phase des rythmes circadiens d'activité neuronale ou motrice. La signalisation par le NPY peut également être classée parmi les signaux non photiques. En effet, si du fait de sa libération par le TGH, le NPY relaie l'information photique, l'IGL d'où partent les projections des neurones NPY, est aussi un centre d'intégration des informations sérotoninergiques en provenance du noyau du Raphé. Tout comme la 5-HT, le NPY va avoir un effet sur la phase des rythmes du NSC seulement durant le jour. « In vitro » l'application de NPY durant le jour subjectif sur des tranches de NSC entraine une avance de phase du rythme d'activité électrique des neurones (Golombek et al 1996; Harrington, & Schak 2000). « In vivo » l'injection de NPY dans le NSC de hamsters maintenus en obscurité constante induit une modification de la phase de l'activité locomotrice (Albers, & Ferris 1984). L'injection faite en début de jour subjectif induit un retard de phase alors qu'une injection faite en fin de jour induit une avance. Enfin, l'effet du NPY durant le jour induit une réduction des niveaux d'expression de Per1 et Per2 (Fukuhara et al 2001; Maywood et al 2002).

### 2.3. Les oscillateurs périphériques

#### 2.3.1. Découverte et démonstration de l'existence des oscillateurs périphériques

Les gènes de l'horloge, acteurs de l'oscillateur moléculaire qui génère l'expression circadienne de nombreux gènes cibles, sont exprimés et rythment dans des cellules autres que celles du NSC. Les fibroblastes de rat en culture présentent une rythmicité d'expression des gènes de l'horloge Per1, Per2, Rer-erba, Dbp après un choc par du sérum (Balsalobre et al 1998). Des souris exprimant dans tous les tissus la luciférase sous le contrôle du promoteur Per1 ou Per2 ont permis de montrer que les tissus périphériques comme le foie, les poumons et les muscles squelettiques expriment des rythmes pouvant durer plusieurs jours mais qui s'atténuent plus rapidement que ceux exprimés dans des tranches organotypiques de NSC (Yamazaki et al 2000; Yoo et al 2004) . Une lignée cellulaire de fibroblastes de souris exprime aussi un rythme circadien des gènes de l'horloge avec un délai entre le pic d'expression de l'ARNm et celui de la protéine correspondante similaire à celui observé dans l'horloge centrale (Yagita et al 2001). L'étude d'un grand nombre de tissus périphériques a permis de mettre en évidence que selon le tissu étudié, le rythme des gènes de l'horloge peut présenter des différences de période et de phase (Yoo et al 2004).

Globalement, l'expression robuste de gènes de l'oscillateur moléculaire a pu être observée dans de nombreux explants de tissus (cœur, poumons, foie, estomac, rate, reins, pancréas, muscles squelettiques, glande tyroïde, glande surrénale) (Yamazaki et al 2000; Yamamoto et al 2004), ainsi que dans plusieurs lignées cellulaires comme les lignées de fibroblastes de souris NIH3-T3 (Akashi, & Nishida 2000) et de rat Rat-1 (Balsalobre et al 1998), la lignée de cellules osseuses humaines U2OS (Hughes et al 2009), la lignée de cellules hépatiques de rat H35 (Balsalobre et al 1998), la lignée hypophysaire de rat GH4C1 (Becquet et al 2014) et la lignée de cellules épithéliales mammaires HC11 (Metz et al 2006). Il est intéressant de noter que les testicules et le thymus sont les seuls tissus où les gènes de l'horloge bien qu'exprimés ne présentent pas de profil circadien d'expression (Morse et al 2003; Alvarez, & Sehgal 2005; Liu et al 2007) . La cause de cette arythmie dans ces tissus reste indéterminée, une explication proposée étant que l'oscillateur circadien ne fonctionne correctement que dans les cellules matures. Cela expliquerait l'absence de rythmicité circadienne dans le

thymus et les testicules qui se composent majoritairement de cellules en cours de maturation et de différenciation (Dibner et al 2010).

### 2.3.2. Robustesse des oscillateurs périphériques

Chez les animaux homéothermes comme les mammifères, la température peut fluctuer de manière importante entre les différents organes. Les organes internes comme le foie ou les reins seront mieux protégés de ces variations que des organes plus exposés comme la peau ou les muqueuses. Sachant que même de petites altérations de la période du rythme circadien peuvent avoir un fort impact sur la phase de ces rythmes, la résistance à la température et la robustesse des oscillateurs périphériques est un critère important dans le maintien d'une rythmicité cohérente dans l'organisme. Les cultures cellulaires de fibroblastes de souris ou de rats sont le modèle cellulaire d'oscillateur périphérique le plus étudié dans les phénomènes de compensation de la température. Un changement de la température d'incubation des fibroblastes n'abolit pas le fonctionnement de leur oscillateur circadien et l'impact des changements de température est compensé pour des variations de 28,5 à 36,5°C (Tsuchiya et al 2003; Izumo et al 2003; Dibner et al 2009) . De manière étonnante, des phénomènes de surcompensation ont même été observés avec une légère accélération du rythme suite à une diminution de la température d'incubation (Izumo et al 2003). Dans des fibroblastes n'exprimant pas le gène Per1 cette surcompensation n'est pas observée suggérant un rôle important de Per1 dans la résistance aux variations de température (Dibner et al 2009).

Les fibroblastes présentent à l'échelle de la cellule isolée une oscillation rythmique continue durant la durée de vie de la cellule. Dans les fibroblastes cette rythmicité persiste même durant la division cellulaire et la phase d'oscillation de la cellule mère est fidèlement retrouvée chez la cellule fille (Nagoshi et al 2004). Ces études indiquent que les oscillateurs périphériques possèdent une rythmicité circadienne extrêmement robuste.

### 2.4. Les mécanismes de synchronisation interne du système circadien

#### 2.4.1 Voies anatomiques de sortie du NSC

Depuis qu'il est apparu clairement que le NSC joue un rôle majeur dans la rythmicité des fonctions circadiennes, de nombreuses recherches se sont intéressées aux voies efférentes du NSC qui permettent la synchronisation des oscillateurs périphériques par le NSC, encore dénommée synchronisation interne du système circadien.

En utilisant des techniques d'autoradiographie antérograde (Swanson, & Cowan 1975; Berk, & Finkelstein 1981; Stephan et al 1981), des traceurs antérogrades (Watts et al 1987; Watts, & Swanson 1987), ou par l'étude anatomique de la dégénération des fibres suite à une lésion du NSC (Stephan et al 1981) les cibles des projections neuronales du NSC ont pu être déterminées. Ces études ont permis d'établir que les projections efférentes du NSC sont au final peu nombreuses et touchent majoritairement des structures de l'hypothalamus et du thalamus.

Dans l'hypothalamus, la voie efférente majeure projette sur la région périsuprachiasmatique et dans la zone sous-paraventriculaire (zSPV). Ces régions sont innervées par les neurones à AVP et à VIP en provenance respectivement de la partie dorso-médiane et ventro-latérale du NSC. Un groupe plus petit de fibres projette vers la région ventrale et latérale du noyau ventromédian de l'hypothalamus. Une autre voie va innerver le noyau parataenial, la partie rostrale du noyau paraventriculaire du thalamus, et de façon moindre, le noyau du lit de la strie terminale. Enfin, le NSC envoie aussi des projections vers la partie ventrale de l'aire préoptique, le septum latéral, le feuillet géniculé latéral ventral ainsi que les noyaux du raphé.

### 2.4.2 Synchronisation des oscillateurs périphériques

L'adaptation des oscillateurs périphériques aux changements de phase de l'environnement a été étudiée à partir de tissus provenant de souris exprimant la luciférase sous le contrôle du promoteur de Per1. En réponse à une avance ou un retard de phase de la photopériode de l'environnement, il a été observé que le changement de rythme s'opère plus vite dans le NSC que dans les tissus périphériques (Yamazaki et al 2000).

La lésion du NSC dans des souris PERIOD2::LUCIFERASE n'induit pas la perte de la rythmicité dans les tissus périphériques. Cependant l'absence du NSC cause une désynchronisation des tissus périphériques entre eux et une plus grande variabilité des rythmes circadiens entre les animaux (Yoo et al 2004). En effet un NSC fonctionnel est nécessaire pour maintenir la cohérence de phase entre les tissus mais aussi entre les cellules d'un même tissu rythmique (Guo et al 2006). Contrairement au NSC dont la robustesse de la rythmicité dépend fortement du réseau formé entre ses neurones, les tissus périphériques présentent plutôt des oscillateurs cellulaires autonomes robustes mais avec une faible capacité de synchronisation (Nagoshi et al 2004). En effet l'enregistrement de la rythmicité de Bmal1 dans des fibroblastes isolés montre que chaque cellule individuellement présente des oscillations circadiennes robustes d'amplitude constante. Cependant même après synchronisation de l'ensemble des fibroblastes isolés, la phase des différentes cellules devient rapidement très variable (Welsh et al 2004).

Les signaux en provenance du NSC sont alors cruciaux au maintien de la synchronisation des cellules d'un même tissu. La nature de ces signaux varie selon les tissus et peut impliquer des voies directes de contact par le NSC. Cependant le nombre de structures contactées directement par le NSC est limité et ces structures sont toutes localisées dans l'hypothalamus et le thalamus. Ce sont donc majoritairement des voies indirectes de nature humorale, hormonale ou nerveuse qui vont permettre la synchronisation des oscillateurs périphériques sur la phase du NSC (Buijs, & Kalsbeek 2001).

L'implication de facteurs humoraux dans la synchronisation des oscillateurs périphériques par le NSC a été suggérée dans les années 90. Chez des hamsters rendus arythmiques par lésion du NSC, la rythmicité circadienne de l'activité locomotrice a pu être restaurée par l'implantation d'hypothalamus antérieur fœtal contenant le NSC. Dans ces expériences, les greffons étaient placés dans des capsules semi-perméables permettant la diffusion de facteurs humoraux mais pas la croissance de projections axonales (Ralph et al 1990; Silver et al 1996). Plus récemment, il a été montré que la synchronisation de fibroblastes en culture peut être améliorée par leur co-culture avec des neurones du NSC (Farnell et al 2011). Ces études montrent que les facteurs humoraux sécrétés par le NSC jouent un rôle majeur dans la régulation de la rythmicité.

Certains rythmes en revanche ne sont pas synchronisés par des facteurs humoraux émis par le NSC. C'est le cas par exemple du rythme de sécrétion de la corticostérone et de la mélatonine qui chez le hamster dont le NSC a été lésé, ne sont pas restaurés par l'implantation de greffon contenant le NSC placé dans des capsules semi-perméables (Silver et al 1996). La restauration de ces rythmes hormonaux nécessite la formation entre le NSC et les oscillateurs endocriniens concernés, de voies neuronales de connexion qui sont indispensables à la transmission des signaux de synchronisation. Pour synchroniser la phase du rythme de sécrétion de la corticostérone par les glandes surrénales, les connections nerveuses que le NSC établit avec le noyau paraventriculaire de l'hypothalamus sont indispensables et permettent aux glandes surrénales de recevoir l'information via la voie hypothalamo-hypophysaire.

### 3. L'hypophyse

### 3.1. Structure et fonctions de l'hypophyse :

L'hypophyse est une des glandes endocrines majeures de notre système endocrinien. Elle se situe dans une cavité osseuse, la fosse hypophysaire, à la base du crane et est protégée par la selle turcique. Elle est connectée à l'hypothalamus dont elle reçoit des afférences par la tige pituitaire. L'hypophyse se divise en deux lobes. Dans la partie antérieure, l'adénohypophyse dérive du plafond de la cavité buccale et dans la partie postérieure, la neurohypophyse est une extension de l'hypothalamus. La neurohypophyse est composée d'une agglomération d'axones de neurones dont les corps cellulaires sont dans l'hypothalamus. Ces axones sont au contact de capillaires sanguins dans lesquels ils vont sécréter les hormones vasopressine et ocytocine. L'adénohypophyse elle, se compose de cinq types cellulaires sécrétant une ou plusieurs hormones : le sous type somatotrope qui sécrète l'hormone de croissance (GH), lactotrope qui sécrète la prolactine (PRL), gonadotrope qui sécrète l'hormone folliculostimulante (FSH) et l'hormone lutéinisante (LH), corticotrope qui sécrète l'hormone adrénocorticotrope (ACTH), thyréotrope qui sécrète la thyréostimuline (TSH). D'autres hormones comme la mélano-stimuline (MSH) et les endorphines obtenues par clivage de l'ACTH sont également sécrétées par l'adénohypophyse. Ces hormones vont être sécrétées dans les capillaires veineux qui irriguent l'hypophyse pour être diffusées dans la circulation générale. La sécrétion de ces hormones est fortement régulée par les signaux en provenance de l'hypothalamus. L'hypothalamus sécrète des neurohormones dans le système porte hypothalamo-hypophysaire qui se situe entre les deux structures, permettant ainsi de communiquer avec les cellules de l'hypophyse. L'hypophyse en sécrétant ces nombreuses hormones dans la circulation sanguine générale va réguler les nombreux tissus qui expriment les récepteurs de ces hormones. Ces récepteurs sont exprimés dans la grande majorité des tissus de l'organisme ainsi que dans d'autres glandes sécrétrices comme la thyroïde, les glandes surrénales, les ovaires et testicules ou encore la glande mammaire. En communiquant avec ces autres glandes, l'hypophyse va moduler la sécrétion hormonale de ces glandes périphériques. La plupart des hormones sécrétées en périphérie régulent en retour la sécrétion

hormonale de l'axe hypothalamo-hypophysaire formant ainsi des boucles d'autorégulation de la sécrétion hormonale.

Par la sécrétion de ses différentes hormones, l'hypophyse participe à la régulation de fonctions physiologiques majeures dans les organismes. L'hormone de croissance GH joue un rôle majeur dans la croissance tissulaire et la régénération (Nussey, & Whitehead 2001). L'hormone thyroïdienne TH qui contrôle la sécrétion des hormones T3 et T4 par la glande thyroïde est ainsi impliquée dans la croissance, la différenciation, le développement et le fonctionnement du métabolisme (Song et al 2011). L'hormone ACTH qui permet la sécrétion des glucocorticoïdes par les glandes surrénales joue une rôle principalement dans la régulation de la réponse au stress mais aussi dans la digestion, le stockage des ressources énergétiques et les émotions (Smith, & Vale 2006). Les hormones sexuelles LH et FSH vont majoritairement contrôler les gonades et participer à la maturation sexuelle et les processus relatifs à la reproduction (Nussey, & Whitehead 2001). Enfin, la prolactine (PRL) est surtout connue pour son rôle dans la stimulation de la production de lait dans la glande mammaire au cours de la lactation (Grattan, & Kokay 2008).

### 3.2. Rythmicité de l'hypophyse

On sait depuis de nombreuses années que les différentes fonctions hormonales de l'hypophyse présentent une rythmicité circadienne. Cependant ce n'est que plus tardivement que l'expression et la sécrétion rythmique des hormones hypophysaires ont pu être montrées (ACTH (Girotti et al 2009), GH (Czeisler, & Klerman 1999), PRL (Dunn et al 1972), LH (Christian et al 2005), TSH (Kalsbeek et al 2000)). Ces hormones présentent non seulement une rythmicité circadienne mais aussi ultradienne et pour certaines d'entres-elles infradienne. Ainsi est diffusée dans tout l'organisme une grande diversité de signaux temporels endocriniens dont l'impact direct sur le fonctionnement physiologique n'est pas encore totalement explicité.

Il est important de noter que dans l'hypophyse, d'autres gènes que les gènes hormonaux sont soumis à une expression circadienne. Pour les identifier, une étude génomique a été réalisée avec la technologie des puces à ADN sur des hypophyses prélevées toutes les heures durant deux jours en condition d'obscurité constante. Suite à cette étude il a été évalué que 400 gènes environ présentent une expression rythmique circadienne dans l'hypophyse (Hughes et al 2007; Hughes et al 2010).

### 3.3. L'hypophyse : un oscillateur autonome

Dans des explants hypophysaires en provenance de modèles murins exprimant la luciférase sous le contrôle du promoteur de Per1, Per2 ou de Bmal1, il a été montré que la luminescence de la luciférase présente une rythmicité circadienne qui persiste durant plusieurs jours (Abe et al 2002; Yoo et al 2004; Chu et al 2013) . Dans des cultures primaires de cellules hypophysaires de rat il a aussi été observé l'expression circadienne du transcrit de Bmal1 (Becquet et al 2014). L'expression circadienne des transcrits de Bmal1, Clock, Per1/2 et Cry1 a aussi été observée dans l'hypophyse des macaques rhésus (Sitzmann et al 2010). Chez la souris il a été montré que cette rythmicité des gènes de l'horloge peut persister quelques jours après une lésion du NSC (Yoo et al 2004). Enfin, l'enregistrement par caméra ultrasensible CDD des variations de la luminescence dans l'hypophyse des souris Per2::Luc, montre que la grande majorité des cellules hypophysaires présentent une luminescence rythmique (Chu et al 2013). Cette étude montre ainsi que chacune des cellules de l'hypophyse contient un oscillateur circadien.

### 3.4. La lignée cellulaire hypophysaire GH4C1

Afin d'étudier les mécanismes moléculaires impliqués dans la rythmicité circadienne hypophysaire, un modèle de lignée cellulaire hypophysaire a été utilisé pour un grand nombre d'expériences présentées dans cette thèse. La lignée hypophysaire GH4C1 est une lignée somatolactotrope produisant les hormones GH et prolactine. Ces cellules ont été pour la première fois produites en mars 1965 et sont issues de tumeurs hypophysaires (MtT/W5) prélevées chez des femelles rat wistar (Tashjian et al 1968). La lignée GH4C1 constitue un modèle de cellules hypophysaire qui a permis principalement l'étude des mécanismes de sécrétion des hormones GH et PRL, de la signalisation neuro-hormonale dans l'hypophyse ainsi que des propriétés électrophysiologiques des cellules sécrétrices (Ooi et al 2004).


D'après Becquet et al., 2014

# Figure 5: Rythme d'expression des gènes de l'horloge et des gènes hormonaux dans la lignée cellulaire GH4C1

Les cellules GH4C1 ont été synchronisées par un changement de milieu puis arrêtées toutes les 4h pendant 30 à 36h. A chaque temps, les ARNs totaux ont été préparés puis analysés par RT-qPCR. Lorsque les valeurs expérimentales s'ajustent à une équation cosinor, dans laquelle la valeur de la période est fixée à 24h, avec un R<sup>2</sup>>0,55, les variations observées sont considérées comme circadiennes. Plusieurs des transcrits de l'oscillateur moléculaire sont exprimés avec une rythmicité circadienne (A), démontrant ainsi la présence d'un oscillateur dans les GH4C1. De plus, les transcrits des deux hormones exprimées dans ce modèle cellulaire hypophysaire, Gh et Prl, ont aussi une expression circadienne (B), ce qui démontre la fonctionnalité de cet oscillateur. La lignée GH4C1 est un modèle intéressant pour l'étude des mécanismes moléculaires impliqués dans les rythmes circadiens car tout comme les explants hypophysaires elle exprime des rythmes circadiens endogènes. En effet, les cellules GH4C1 peuvent être synchronisées par un changement de leur milieu de culture et présentent alors une rythmicité circadienne d'expression des transcrits de Bmal1, Per2, Rerv-erb $\alpha$  et Ror $\gamma$  (Becquet et al 2014) (Fig5). L'expression dans ces cellules du gène rapporteur luciférase sous le contrôle du promoteur de Per1 ou Npas2 montre une activation circadienne de ces promoteurs. De plus, l'expression d'un dominant négatif de Bmal1 (Hosoda et al 2004) dans ces cellules induit une perte de la rythmicité circadienne des autres principaux gènes de l'oscillateur moléculaire circadien (Becquet et al 2014). Enfin, l'oscillateur moléculaire dans cette lignée cellulaire est fonctionnel, comme en atteste l'expression rythmique de nombreux gènes. Parmi ces gènes se trouvent ceux des deux hormones sécrétées par les cellules GH4C1, GH et PRL (Guillaumond et al 2011; Becquet et al 2014) (Fig5).

## 4. Les rythmes d'expression génique

### 4.1. Importance quantitative des transcriptomes cellulaires

Les fonctions circadiennes des organismes reposent sur l'expression rythmique des gènes générée par l'oscillateur moléculaire circadien. Plusieurs études se sont attelées à déterminer l'importance du transcriptome rythmique dans les différents organismes. Par la technologie des puces à ADN, ou « microarray », il a pu être montré qu'un nombre important de gènes est exprimé de manière circadienne dans l'organisme. Par exemple, dans le NSC de souris 5 à 10 % des gènes présentent une expression circadienne (Panda et al 2002; Ueda et al 2002) . Ce même type d'étude réalisée chez différentes espèces (souris, drosophile, fungis, plantes) et pour les mammifères dans différents tissus (cerveau, NSC, foie, rein, cœur etc...) a permis d'évaluer qu'entre 8% et 10% des gènes sont rythmiques en moyenne dans un tissu (Storch et al 2002; Panda et al 2002; Duffield 2003; Vollmers et al 2009; Lowrey, & Takahashi 2004; Hughes et al 2009) . La comparaison des listes de gènes rythmiques dans les différents tissus a mis en évidence la forte spécificité tissulaire de la rythmicité d'un gène que la

comparaison soit réalisée entre le foie et le coeur (Storch et al 2002), le foie et le cerveau (Panda et al 2002) ou encore le foie et les muscles (Miller et al 2007). En effet moins de 5% des gènes rythmiques sont communs à deux oscillateurs. Plus récemment, le développement des techniques de séquençage haut débit des ARNs (RNA-seq) a amélioré la résolution des études et a permis de reconsidérer à la hausse l'importance du transcriptome rythmique dans les organismes. Ainsi chez la souris, l'étude par RNA-seq du transcriptome de douze organes montre que 43% des gènes codants pour des protéines présentent une rythmicité circadienne dans l'un des organes (Zhang et al 2014). De manière intéressante, plus de 1000 ARNs ne codant pas pour des protéines ont aussi été trouvés exprimés avec une rythmicité circadienne (Zhang et al 2014). La même observation est faite chez la drosophile où 10% des gènes exprimés avec une rythmicité circadienne dans le cerveau sont des ARNs non codants (Hughes et al 2012). Enfin, face à la multiplication des études génomiques du transcriptome rythmique, des bases de données ont été développées afin de regrouper les résultats obtenus par différentes équipes dans différents modèles. L'interface CirGRDB est la base de données la plus récemment développée (Li et al 2017). Elle regroupe plus de 4000 études génomiques du transcriptome circadien, réalisées dans des tissus ou des cultures cellulaires d'humain ou de souris.

A l'origine de la rythmicité de ces gènes se trouve l'oscillateur moléculaire. Les gènes de l'horloge qui constituent l'oscillateur vont induire l'expression rythmique des gènes contrôlés par l'horloge selon différents mécanismes. Ainsi les gènes contrôlés par l'horloge vont soit être des cibles directes des gènes de l'horloge, ou vont être des cibles indirectes dont la rythmicité est induite par des facteurs intermédiaires eux mêmes sous le contrôle des gènes de l'horloge. Les mécanismes de régulation directe impliquent un niveau de contrôle transcriptionnel alors que les mécanismes de régulation indirecte impliqueront aussi bien des niveaux de contrôle transcriptionnels que post-transcriptionnels.

## 4.2. Mécanismes transcriptionnels de contrôle

Au niveau transcriptionnel, le contrôle direct de la rythmicité des gènes par les gènes de l'oscillateur moléculaire nécessite des séquences consensus de reconnaissance par

les gènes de l'oscillateur moléculaire. L'hétérodimère BMAL1:CLOCK va se fixer sur des séquences de type E-box située en cis- dans le promoteur des gènes cibles de l'horloge circadienne (Gekakis et al 1998). En ce fixant sur cette séquence, BMAL1:CLOCK va induire la transcription rythmique des gènes. De manière similaire la protéine REV-ERB $\alpha$  va se fixer sur la séquence consensus RORE présente elle aussi dans le promoteur de certains gènes (Yin et al 2007). Du fait de la rythmicité d'expression de REV-ERB $\alpha$ , la transcription des gènes présentant une séquence RORE sera elle aussi rythmique.

En dehors de ces séquences qui permettent une régulation directe par les gènes de l'horloge, plusieurs autres mécanismes peuvent entraîner de manière indirecte la transcription rythmique de gènes.

La protéine de liaison à l'albumine D (DBP) est une voie indirecte de contrôle de la rythmicité de transcription (Okamura et al 2002). La protéine DBP dont le promoteur présente une séquence E-Box a de ce fait une expression rythmique sous le contrôle direct de BMAL1:CLOCK. La protéine DBP va à son tour se fixer sur des séquences D-box consensus présentes dans le promoteur de certains gènes et contrôler de manière rythmique leur transcription. La protéine DBP est donc un facteur de transcription intermédiaire entre les gènes de l'horloge et les gènes cibles contrôlés par l'oscillateur. Ce rôle peut aussi être attribué à d'autres facteurs de transcription dont la rythmicité d'expression peut être régulé via un ou plusieurs facteurs relais du même ordre que la protéine DBP.

Enfin, l'oscillateur circadien peut contrôler encore plus indirectement la transcription rythmique de gène en étant responsable de l'expression rythmique de co-facteurs de transcriptions qui vont à leur tour imposer leur rythme aux gènes dont ils co-régulent la transcription. Ainsi, l'activation circadienne du promoteur de la prolactine dans les cellules hypophysaires va dépendre de l'expression circadienne des co-facteurs de transcription NONO et SFPQ qui s'associent à des facteurs de transcription qui eux ne sont pas rythmiques (Guillaumond et al 2011).

### 4.3. Mécanismes post-trancriptionnels

La caractérisation des mécanismes moléculaires à la base de la genèse des oscillations circadiennes a constitué une avancée considérable dans le domaine de la chronobiologie. On sait aujourd'hui que l'oscillation circadienne repose sur un mécanisme intracellulaire, très conservé d'une espèce à l'autre, impliquant un groupe de gènes de l'horloge et leurs protéines respectives (Ko, & Takahashi 2006). Le modèle proposé actuellement implique différentes boucles de rétroaction transcriptionnelle positives et négatives. Un des axes majeurs développés en chronobiologie a trait à la recherche des mécanismes moléculaires qui vont permettre l'expression rythmique des gènes. La nature de ces mécanismes est très diverse puisque le contrôle rythmique peut s'exercer au niveau de la transcription des gènes (Ripperger, & Brown 2010), de leur régulation post-transcriptionnelle (Petrillo et al 2011), et même au niveau posttraductionnel, certaines protéines ayant en effet une expression ou une activité rythmique sans rythmicité d'expression de leurs ARNm (Jouffe et al 2013). En effet, des études protéomiques réalisées dans le foie de souris montrent que presque la moitié des protéines dont l'expression est rythmique résulte de la traduction d'ARNm arythmiques (Reddy et al 2006; Mauvoisin et al 2014).

Si de nombreuses études dont celles réalisées dans l'équipe (Guillaumond et al 2011; Becquet et al 2014), se sont focalisées sur les régulations transcriptionnelles rythmiques, on sait aujourd'hui que la rythmicité des ARNm ne repose pas exclusivement sur une transcription rythmique (Koike et al 2012; Menet et al 2012; Le Martelot et al 2012; Lück et al 2014; Atger et al 2015).

L'un des premiers exemples de régulation post-transcriptionnelle de la rythmicité circadienne a été mis en évidence chez le dinoflagellé Gonyaulax Polyedra. Cette espèce synthétise la protéine luciférase avec une rythmicité circadienne. Pourtant l'expression de l'ARNm de la protéine luciférase est arythmique (Morse et al 1989). Il existe donc des mécanismes intervenant pendant ou après la traduction qui induisent l'expression circadienne de la protéine luciférase. Un autre exemple de l'implication de mécanismes post-transcriptionnels dans la rythmicité circadienne a été observé chez la drosophile. Les auteurs de cette étude soulignent une différence entre la rythmicité d'activation de la transcription et celle de l'expression de l'ARNm pour deux gènes principaux de l'oscillateur moléculaire de la drosophile, Per et Cry-1 (So, & Rosbash 1997) et

suggèrent que des mécanismes post-transcriptionnels s'associent à des mécanismes transcriptionnels pour contrôler le rythme d'expression de ces gènes.

Plusieurs années plus tard, le développement des technologies de séquencage du génome a permis non seulement de quantifier le nombre de gènes par tissu dont l'expression est rythmique mais aussi d'étudier le rôle des régulations posttranscriptionnelles dans l'expression rythmique des gènes. Deux premières études réalisées sur du tissu hépatique de souris ont montré qu'environ 70% des ARNm exprimés de manière circadienne sont contrôlés par des mécanismes posttranscriptionnels (Koike et al 2012; Menet et al 2012). Plus tard, cette évaluation sera pondérée par l'utilisation d'une approche alternative d'analyse de ces mêmes données brutes. Cette dernière analyse tient compte des phénomènes de dégradation des ARNm qui, en modifiant la demi-vie des ARN, influencent leur phase d'expression. Avec cette analyse, la proportion des ARNm dont la rythmicité dépend de régulations posttranscriptionnelles est réévaluée à environ 30% Lück 2014. Une autre étude vient corroborer cette dernière estimation quantitative en déterminant, à partir du profil de transcription, d'accumulation et de traduction des ARNs du foie de souris, que les régulations transcriptionnelles sont responsables de l'accumulation rythmique des ARNm dans près de 70% des cas Atger 2015.

Bien que les estimations quantitatives de leur implication diffèrent, il apparaît que les régulations post-transcriptionnelles jouent un rôle majeur dans la rythmicité du transcriptome hépatique.

L'importante contribution des mécanismes post-transcriptionnels a aussi été établie pour d'autres organismes, dont Neurospora, comme en atteste le fait que chez cette espèce, la liste des gènes dont les ARNm présentent un rythme circadien d'expression ne recouvre que très partiellement la liste des gènes dont le promoteur est activé de manière rythmique (Hurley et al 2014). Nous avons été invités à rédiger une revue concernant les mécanismes posttranscriptionnels impliqués dans le contrôle de l'expression rythmique des gènes dans le journal WIREs RNA :

## Circadian processes in the RNA life cycle

Manon Torres, Denis Becquet, Jean-Louis Franc and Anne-Marie François-Bellan WIREs RNA, 2018, sous presse. **DOI:**10.1002/wrna.1467

Dans cette revue, nous présentons l'ensemble des mécanismes décrits chez les mammifères qui en s'exerçant après la transcription des gènes, sont capables d'induire leur expression circadienne. Ces mécanismes concernent majoritairement les différentes étapes de la maturation des ARNm et incluent l'ajout de la coiffe à l'extrémité 5', l'épissage alternatif, les modifications de l'efficacité de l'épissage et les changements de stabilité de l'ARN contrôlés par la longueur de la queue polyadénylée ou l'utilisation de sites de polyadénylation alternatifs. Le transport des ARNs peut également suivre un rythme circadien, avec une rétention nucléaire rythmique induite par le rythme du nombre de paraspeckles dans le noyau des cellules, et une exportation rythmique vers le cytoplasme qui est guidée par des protéines rythmiques agissant comme des cargos. Enfin, la dégradation des ARNs peut également suivre un rythme circadien au travers de l'implication rythmique de miRNAs.

DOI: 10.1002/wrna.1467

#### ADVANCED REVIEW



## Circadian processes in the RNA life cycle

Manon Torres | Denis Becquet 10 | Jean-Louis Franc 10 | Anne-Marie François-Bellan 10

CNRS, CRN2M-UMR7286, Faculté de Médecine Nord, Aix-Marseille Université, Marseille, France

**Correspondence** Email: anne-marie.francois@univ-amu.fr The circadian clock drives daily rhythms of multiple physiological processes, allowing organisms to anticipate and adjust to periodic changes in environmental conditions. These physiological rhythms are associated with robust oscillations in the expression of at least 30% of expressed genes. While the ability for the endogenous timekeeping system to generate a 24-hr cycle is a cell-autonomous mechanism based on negative autoregulatory feedback loops of transcription and translation involving core-clock genes and their protein products, it is now increasingly evident that additional mechanisms also govern the circadian oscillations of clock-controlled genes. Such mechanisms can take place post-transcriptionally during the course of the RNA life cycle. It has been shown that many steps during RNA processing are regulated in a circadian manner, thus contributing to circadian gene expression. These steps include mRNA capping, alternative splicing, changes in splicing efficiency, and changes in RNA stability controlled by the tail length of polyadenylation or the use of alternative polyadenylation sites. RNA transport can also follow a circadian pattern, with a circadian nuclear retention driven by rhythmic expression within the nucleus of particular bodies (the paraspeckles) and circadian export to the cytoplasm driven by rhythmic proteins acting like cargo. Finally, RNA degradation may also follow a circadian pattern through the rhythmic involvement of miRNAs. In this review, we summarize the current knowledge of the post-transcriptional circadian mechanisms known to play a prominent role in shaping circadian gene expression in mammals.

This article is categorized under:

RNA Processing > Splicing Regulation/Alternative Splicing

RNA Processing > RNA Editing and Modification

RNA Export and Localization > Nuclear Export/Import

#### KEYWORDS

circadian rhythms, degradation, editing, nuclear export, miRNA, nuclear retention, paraspeckles, RNA processing, RNA stability, splicing

### **1 | INTRODUCTION**

#### **1.1** | What are circadian rhythms?

Due to Earth's rotation, living organisms must live in an environment that changes daily. Organisms with the ability to anticipate and synchronize themselves with their rhythmic environment have evolutionary-selective advantages, and this has driven the evolution of endogenous clocks across all domains of life (bacteria, fungi, plants, invertebrates, and mammals) (Bell-Pedersen et al., 2005; Harmer, Panda, & Kay, 2001). In rodents, the suprachiasmatic nucleus (SCN) of the hypothalamus consists of ~ 20,000 neurons and functions as the master oscillator of the endogenous clock. However, the molecular

## <sup>2 of 13</sup> WILEY WIREs

mechanism of this endogenous clock is cell autonomous and is present in almost every cell of the body. Within each tissue, circadian oscillator cells must be coupled with each other for circadian organization to emerge at the tissue level. The SCN plays a special role within this multioscillatory system by providing signals that maintain the coupling of oscillators within most tissues. The SCN timing signals are conveyed to peripheral tissues and extra-SCN brain regions by neural networks, hormonal pathways (such as glucocorticoids), and behavioral pathways (such as sleep/wake or food intake rhythms) (Abe et al., 2002; Guilding, Hughes, Brown, Namvar, & Piggins, 2009; Hastings, Maywood, & Reddy, 2008; Menaker, Murphy, & Sellix, 2013; Schibler & Sassone-Corsi, 2002; Yoo et al., 2004). It should be noted that peripheral clocks not only respond to the SCN resetting of signaling but can also be reset by an alteration of the food intake rhythm, which permits the local integration of systemic and nutritional signals (Potter, Cade, Grant, & Hardie, 2016).

The endogenous clock system generates circadian rhythms characterized by a period of approximately 24 hr. These can vary in their amplitude and/or phase. While external cues (inputs) maintain the synchronization of the circadian clock within its environment, under constant conditions, those rhythms are persistent. The circadian clock system controls the rhythmicity of a wide range of processes (outputs), from molecular systems (e.g., transcriptional mechanisms), cellular systems (e.g., the cell cycle (Hong et al., 2014)) physiological level (hormonal secretion (Tonsfeldt & Chappell, 2012), metabolic level (Bailey, Udoh, & Young, 2014)), up to the behavioral level (sleep/wake cycle, cognitive functions, etc.). Therefore, it has become abundantly clear that a functional circadian clock is a fundamental element in health. Numerous studies have shown how a dysfunctional circadian clock can lead to major diseases (neurodegenerative disorders (Musiek, 2015; Videnovic, Lazar, Barker, & Overeem, 2014), obesity and metabolic diseases (Albrecht, 2017), cancer (Kettner, Katchy, & Fu, 2014; Soták, Sumová, & Pácha, 2014), reproductive disorders (Sellix, 2015), and cardiovascular diseases (Takeda & Maemura, 2015)).

#### 1.2 | The molecular circadian oscillator

Within each vertebrate cell, the molecular mechanism of the autonomous oscillator relies on core-clock genes (CGs). Even if CGs vary among species, this molecular mechanism is mostly conserved. Indeed, in most vertebrate species, this mechanism is sustained by interlocked transcriptional-translational feedback loops (TTFL) of activators and repressors that form a canonical positive-feedback and negative-feedback gene network motif (Novák & Tyson, 2008). The TTFL cycle has a period of approximately 24 hr (for reviews, see Cohen & Golden, 2015; Hernando, Romanowski, & Yanovsky, 2017; Hurley, Loros, & Dunlap, 2015; Ozkaya & Rosato, 2012; Takahashi, 2015).

In mammals, BMAL1/CLOCK dimers form the positive arm of the core loop and activate the transcription of the *Per* (*Per1*, *Per2* and *Per3*) and *Cry* (*Cry1* and *Cry2*) genes by binding a specific E-box present in promoter regions. In turn, PER/CRY heterodimers translocate back to the nucleus and repress their own transcription by acting on the BMAL1/CLOCK complex, thus forming the negative arm of the loop and completing the core TTFL of the mammalian circadian oscillator (see Figure 1(a)).

This primary loop is reinforced by an additional interlocking loop involving the retinoic acid receptor-related genes *Rora*, *Ror* $\beta$ , *Ror* $\gamma$ , *RevErba and RevErb* $\beta$ , whose products in turn respectively stimulate and repress the transcription of the *Bmal1* gene through the Retinoic acid receptor-related Orphan Receptor Elements (ROREs) present in the *Bmal1* promoter, further stabilizing the circadian circuitry (Guillaumond, Dardente, Giguère, & Cermakian, 2005; Preitner et al., 2002; Sato et al., 2004) (see Figure 1(a)).

In addition, the molecular circadian oscillator drives the circadian expression of a wide range of clock-controlled genes (CCGs) through E-boxes or ROREs present in their promoter. When E-boxes or ROREs are lacking in CCG promoters, the circadian oscillator can use indirect pathways to control the transcription of CCGs. Those pathways involve gene expression cascades that include transcription factors such as albumin D-box binding protein (DBP) or E4BP4 (also known as NFIL3) as immediate circadian oscillator targets (Panda et al., 2002). A number of CCGs are indeed under the circadian regulation by the way of response elements for some of these factors, such as D-boxes and CRE elements (Ueda et al., 2005). It has also been shown that the circadian oscillator can control CCGs through the circadian control of cotranscriptional factors that relay circadian information to their own targets (Guillaumond et al., 2011) (see Figure 1(a)).

The identification of genes that are downstream of the circadian pacemakers using DNA microarray technology (Duffield, 2003; Lowrey & Takahashi, 2004; Panda et al., 2002; Storch et al., 2002; Vollmers et al., 2009) or high-throughput RNA sequencing (Hughes, Grant, Paquin, Qian, & Nitabach, 2012) has allowed for a quantitative evaluation of the rhythmic transcriptomes of different species (rodents, drosophila, fungi, and plants). In mammals, this study has extended to the observation of different tissues (SCN, liver, heart, kidney, etc.) as well (Zhang, Lahens, Ballance, Hughes, & Hogen-esch, 2014). The recent advent of genomic techniques was instrumental in determining the large size of the rhythmic output network comprising 6%–20% of the expressed genome, depending on the cell type or tissue and environmental conditions (Hughes et al., 2012; Li, Grant, Hogenesch, & Hughes, 2015). Such genome-wide studies performed in various mouse tissues indicated that each tissue expresses its own particular set of circadian-regulated genes, which only partly overlap with

manner

-WILEY WIREs 3 of 13





each other (Miller et al., 2007; Panda et al., 2002; Zhang et al., 2014). It has even been projected that more than half of the mouse protein-coding genome shows circadian oscillation somewhere in the body (Zhang et al., 2014).

#### **1.3** | The significance of post-transcriptional regulation

Until recently, transcription was believed to be the main rhythm-inducing mechanism. However, results obtained in the last several years instead indicate that approximately 25% of circadian messenger RNAs (mRNAs) are regulated by de novo transcription, suggesting that other mechanisms account for the majority of circadian gene expression (Hurley et al., 2014; Koike et al., 2012; Menet, Rodriguez, Abruzzi, & Rosbash, 2012). Indeed, several recent reports have confirmed that many, if not all, steps in the life cycle of a (pre-) mRNA (see Figure 1(b)) may be regulated in a circadian manner, thus contributing to circadian gene expression. These mechanisms were first described in the dinoflagellate *Lingulodinium* (Morse, Milos, Roux, & Hastings, 1989) but have since been reported in plants, fungi, Drosophila, and mammals. Genome-wide studies

## 4 of 13 WILEY WIRES

have only recently made it possible to determine how important post-transcriptional mechanisms are for circadian rhythmicity. The study of the hepatic circadian proteome of the mouse has shown that among the proteins displaying a circadian expression pattern, half lacked a corresponding cycling transcript (Reddy et al., 2006). These results highlight the major importance of the post-transcriptional mechanisms in circadian rhythmicity and have led to many recent studies describing how mRNAs rhythmicity can be achieved through post-transcriptional regulation (see Figure 1(b)). In addition to providing emerging evidence for an important role of post-transcriptional regulation, from splicing, polyadenylation, and mRNA stability to translation and noncoding functions exemplified by microRNAs, these studies have also highlighted that this level of regulation affects virtually all aspects of circadian systems, from the core timing mechanism and input pathways that synchronize the environmental clocks, to the output pathways that control overt rhythmicity.

Post-transcriptional processes begin even while the RNA is still being transcribed and occur in a complex and highly coordinated manner. These processes involve large protein complexes, including many RNA-binding proteins (RBPs), ribonucleases with specific functions, and the extensive machinery that conducts the microRNA-mediated control of mRNA stability and translation to ultimately control gene expression. Among RBPs, some proteins remain stably bound to RNA while others are subject to dynamic exchanges that can in turn have a profound impact on circadian rhythms.

It has been shown that post-transcriptional regulatory mechanisms may be driven by temperature, thus adding another level of complexity to the process of circadian gene expression (Buhr, Yoo, & Takahashi, 2010; Gotic et al., 2016; Gotic & Schibler, 2017). Among these mechanisms, *trans*-acting systems would likely imply RBPs with accumulation and/or activity controlled by temperature. On the other hand, *cis*-acting mechanisms may involve temperature-dependent RNA secondary structures that favor transcript degradation or splicing.

Below, we review the recent advances in research of these various post-transcriptional RNA-based mechanisms, focusing on recently described mechanisms in mammals. However, it should be noticed that works performed in others species, especially in flies, have provided discoveries that may be related to mechanisms discovered in mammals and can be very useful for their comprehension. The group of Michael Rosbash, who was awarded together with two others chronobiologists by the 2017 Nobel Prize in physiology or medicine for his discoveries of molecular mechanisms controlling the circadian rhythm, has done a lot of work on post-transcriptional mechanisms in the clock of the fly, starting from initial observations in the 1990s (Zwiebel, Hardin, Liu, Hall, & Rosbash, 1991), up to mechanical insights in recent years (Bartok, Kyriacou, Levine, Sehgal, & Kadener, 2013; Kadener et al., 2009; Lerner et al., 2015; Rodriguez et al., 2013). The role of alternative splicing and miRNAs in the clock functioning of the fly has been particularly documented (Bartok et al., 2013; Kadener et al., 2009; Lerner et al., 2013).

#### 1.4 | Circadian RNA processing

Every step in RNA processing including 5'UTR regulation by capping-dependent as well as capping-independent mechanisms, splicing events and 3'UTR regulation specifically by poly(A) tail regulation (length or alternative poly[A] site) are under circadian timekeeping control.

#### 1.5 | Circadian 5'UTR capping

Processing of mammalian pre-mRNA begins with the formation of the cap structure in the 5'UTR, a process involving the association of the eukaryotic translation initiation factor 4E (eIF4E) with the 5' cap structure. There are no data showing that the regulation of capping is under circadian control. However, there are reports showing that eIF4E is involved in the control of circadian translation (Cao et al., 2013; Cao et al., 2015; Saraf, Luo, Morris, & Storm, 2014). It has been shown that the CG protein BMAL1 is rhythmically phosphorylated by the ribosomal S6 protein kinase 1 (S6 K1) and interacts upon phosphorylation with the translation machinery through the cap-binding protein complex eIF4E. This process controls cap-dependent translation in a circadian manner (Lipton et al., 2015). Phosphorylated BMAL1 is therefore believed to provide a scaffold for more efficient assembly of the translational machinery that gives rise to a general increase in cap-dependent translation.

#### 1.6 | Circadian 5'UTR and 3'UTR binding

An internal ribosomal entry site (IRES) in the 5'UTR of *Per1* mRNA regulates its translation. IRES-regulated translation allows for the direct recruitment of the 40S ribosome subunit to the vicinity of the initiation codon, and this process is therefore independent of the classical cap-dependent translation machinery. The IRES element of *Per1* mRNA is implicated in the rhythmic expression of the protein by means of its rhythmic interaction with the heterogeneous nuclear ribonucleoprotein Q (hnRNPQ). HnRNPQ is rhythmically phosphorylated and activates *Per1* translation in a circadian manner (Lee et al., 2011; Lee, Kim, Kim, & Kim, 2012). The translation of *Reverba* is also rhythmically regulated through an interaction with an IRES, hnRNPQ and polypyrimidine tract-binding protein (PTB) (Kim, Woo, Lee, Kim, & Kim, 2010). The transcription factor

## WILEY WIREs 5 of 13

Nuclear factor interleukin 3 (NFIL3), which binds to a D-box element residing in the promoter of CGs and contributes to their robust oscillation, contains in its 5'UTR an IRES-mediated translation that contributes to the rhythmic expression of NFIL3 protein through an interaction with the RBP hnRNPA1 (Kim et al., 2017). Within the pineal gland, the rate-limiting enzyme in the melatonin biosynthetic pathway, arylalkylamine-N-Acetyltransferase (AANAT), exhibits circadian expression that is driven by circadian IRES-mediated translation. The circadian *Aanat* IRES activity is mediated by the rhythmic expression of hnRNPQ, which specifically interacts with an adenosine-rich region in the 5'end of the *Aanat* IRES (Kim et al., 2007).

In addition to the 5'-UTR, the 3' UTR of mRNA is known as a hotspot for post-transcriptional gene regulation, controlling important cellular functions such as morphogenesis, cell differentiation, metabolism, cell proliferation, and other processes by controlling mRNA translation, stability, localization, and 3' end processing. Different RBPs have been shown to control CGs or CCGs by binding to sequences located in the 3'UTR that are known to be important for the maturation and stability of pre-RNAs and mRNAs. Whether this circadian control relies on the circadian expression of theses RBPs or on their circadian cellular re-localization remains undetermined. It has been shown that while hnRNPK stabilizes *Per3* mRNA by binding to its 3'UTR, thus increasing the amplitude of circadian mouse *Per3* mRNA oscillation, hnRNPD is a destabilizer that decreases its levels (Kim et al., 2011; Kim et al., 2015; Kwak, Kim, & Kim, 2006). Importantly, hnRNPD has been further implicated in the destabilization of *Cry1* (Woo et al., 2010). Another RBP, hnRNPQ, functions by binding to both the 3' and 5'UTRs of mouse *Per3* mRNA, enhancing its decay(Kim et al., 2015). The cooperative function of the 5' and 3'UTRs is necessary for the critical role played by HnRNPQ in the maintenance of the circadian oscillation of mouse *Per3* (Kim et al., 2011).

LARK, a mouse RBP, is rhythmically expressed in the SCN and binds to a cis element in the 3'UTR of *Per1*, leading to an increase in *Per1* translation (Kojima et al., 2007). An isoform of the U2-auxiliary-factor 26 (U2AF26), resulting from circadian-based alternative splicing, destabilizes PER1 (Preußner et al., 2014). PTB, another RBP and a member of the hnRNP family (hnRNPI, gene name *Ptbtp1*), destabilizes *Per2* mRNA by binding to its 3'UTR (Woo et al., 2009).

#### 1.7 | Circadian Splicing

#### 1.7.1 | Alternative splicing

Another step of gene expression that could fine-tune the circadian clock is pre-mRNA splicing, a mechanism catalyzed by the spliceosome complex in which introns are removed and exons are joined together to form mature mRNAs. In many cases, the splicing process can create alternative mRNAs from the same precursor by varying the selection of the included/ excluded regions. As a rule, this mechanism, known as alternative splicing (AS), can affect the coding regions of a gene, resulting in the production of multiple protein isoforms or can affect non-coding regions, giving rise to transcript variants that harbor different regulatory elements that influence the translational efficiency or stability of the mRNA (Chen & Manley, 2009). AS events have mostly been studied in the context of the temperature control of circadian system functioning.

Alternative splicing may be regulated in a circadian manner and may control different aspects of the circadian clock. A systematic study of mammalian circadian alternative splicing recently performed using a microarray-based approach in mouse liver has shown that many exons are alternatively spliced in a circadian manner. Furthermore, in this microarray analysis, several RBPs that may be involved in the regulation of circadian RNA processing were shown to display a circadian mRNA expression pattern (McGlincy et al., 2012). Proteins such as SFPQ and NONO, which regulate alternative splicing in different contexts, are known to be involved in clock function (Duong, Robles, Knutti, & Weitz, 2011; Guillaumond et al., 2011; Heyd & Lynch, 2010; Kowalska et al., 2012). Preußner et al. (Preußner et al., 2014) have established the first functional connection between alternative splicing and the circadian clock in mammals. These authors described a robust circadian and light-inducible splicing switch within exons 6 and 7 that changes the reading frame of the mouse mRNA encoding U2AF26. Interestingly, this isoform modifies the C terminus of the protein, adding a new domain that shows homology to parts of the Drosophila clock regulator called TIMELESS, which is a central component of the fly circadian clock. As it does in TIMELESS, this domain allows the new U2AF26 variant to interact with PER1, thereby regulating its stability (as mentioned above) and thus its core-clock function (Preußner et al., 2014). Mice deficient in U2AF26 show nearly arrhythmic PER1 protein levels and broad defects in circadian mRNA expression in peripheral clocks. At the behavioral level, these mice display increased phase advance adaptations following experimental jet lag, suggesting that light-induced U2af26 alternative splicing allows for the control of PER1 induction and the avoidance of abnormal changes in the circadian clock that may occur under light/dark conditions. Taken together, these findings provide evidence for a functional role of U2af26 alternative splicing in controlling the circadian clock (Preußner et al., 2014). More recently the same group shows that body temperature cycles drive rhythmic SR protein phosphorylation to control rhythmic alternative splicing (Preußner et al., 2017). A temperature change of 1°C leads to the splicing switch of a group of RNAs including U2af26 and TATAbox binding protein (Tbp). Since in the later case, the alternative splicing results in rhythmic TBP protein levels with implications for global gene expression, the splicing-based thermometer described by these authors may have a broad impact on circadian gene expression.

The murine melanopsin gene (*Opn4*) provides another example of functional alternative isoform expression. This gene is indeed alternatively spliced, producing two distinct isoforms, a short and a long isoform, which differ only in their C terminus tails. Both form fully functional photopigments (Pires et al., 2009). These different isoforms of *Opn4* have been shown to mediate different behavioral responses to light (Jagannath et al., 2015). Whether these two isoforms are rhythmically produced remains to be determined.

The identification of circadian exons within *Clock*, *Npas2* and *Reverbα* suggests the existence of a feedback loop wherein the clock regulates alternative splicing and, in turn, is regulated by alternative splicing (McGlincy et al., 2012). Two other core components of the mammalian circadian clock, *Cry1* and *Per2*, was also shown alternatively spliced in a circadian manner (Avitabile et al., 2014; Pembroke, Babbs, Davies, Ponting, & Oliver, 2015). However, the physiological role of these isoforms remains to be elucidated.

#### 1.7.2 | Splicing efficiency

Beside circadian AS that contributes to circadian gene expression, Gotic et al. have recently reported evidence regarding temperature-dependent circadian splicing efficiency (Gotic et al., 2016; Gotic & Schibler, 2017). The authors have shown that the primary transcripts of the cold-inducible RBP, *Cirbp* pre-mRNAs, while synthesized at similar rates whatever the temperature level, are more quickly processed into mature mRNAs at lower temperatures than at higher temperatures. This higher splicing efficiency at lower temperatures results in an increased accumulation of mature *Cirbp* mRNA and protein. Furthermore, the authors observed that this mechanism, in addition to controlling mRNA stability, appears to be a wide-spread feature in the temperature-dependent regulation of circadian mammalian gene expression. They then proposed that the controlled fraction of pre-mRNA processed into mature functional mRNA is a novel temperature-dependent post-transcriptional mechanism that modulates gene output. In contrast to the well-studied process of alternative splicing, the mechanism they describe governs the temperature-dependent accumulation of one and the same mRNA species (Gotic et al., 2016; Gotic & Schibler, 2017).

#### **1.8** | Poly(A) tail length and Alternative Polyadenylation

One key feature of eukaryotic mRNAs is the co-transcriptionally added 3'-poly(A) tail, which is important for their nuclear export, translation and stability. More recently, alternative cleavage and polyadenylation (APA) has been found to represent an important layer of post-transcriptional gene regulation, which in turn can influence RNA fate and/or regulate the protein output quantitatively or qualitatively, thereby driving important cellular programs (Jalkanen, Coleman, & Wilusz, 2014).

The length of the poly(A) tail provides the first level of post-transcriptional control (Weill, Belloc, Bava, & Méndez, 2012), with long tails having a stabilizing effect in the bulk of mRNAs. In contrast, (A)-tail shortening generally initiates the degradation of mRNAs. Changes in poly(A) tails have been implicated in a number of physiological and pathological processes. Poly(A) tail length may be under circadian clock control. It has been shown that *nocturnin* (ccnr4l), a deadenylase that degrades the tails of specific mRNA targets, is rhythmically expressed in various murine tissues (Garbarino-Pico et al., 2007; Wang et al., 2001). A daily variation in the poly(A) tail length has also been described for two mRNA species, vasopressin and Fabp7 (Gerstner et al., 2012; Robinson, Frim, Schwartz, & Majzoub, 1988). Another study (Kojima, Sher-Chen, & Green, 2012) extended this observation to a large number of additional mRNAs. In this study, the authors showed that 2.3% of the mRNAs in the mouse displayed rhythmic modification of their poly(A) tail length. These mRNAs can be separated into three groups: group I in which the pre-RNA and mRNA are rhythmic, group II in which the pre-RNA is rhythmic while the mRNA is not, and group III in which neither the pre-RNA nor the mRNA displays rhythmicity. The length of the poly(A) tail may be regulated in the nucleus, through transcription-coupled poly(A) tail synthesis, or in the cytoplasm, through cytoplasmic polyadenylation element-binding protein (CPEB)-mediated cytoplasmic polyadenylation. The regulation of polyadenylation occurs in the nucleus for the two first groups, whereas this regulation takes place in the cytoplasm for group III. For this latter group, the proteins are rhythmically expressed while the steady-state mRNA levels are not. There is a strong correlation between the poly(A) tail length rhythmicity and protein rhythms, as the time of peak tail length precedes the peak in protein expression by several hours (Kojima et al., 2012). Cytoplasmic circadian polyadenylation occurring in the liver may be mediated by CPEB2 and CPEB4, which exhibit a circadian expression pattern (Kojima et al., 2012).

In addition to the length of the poly(A) tail, the site of polyadenylation within the 3'UTR region of mRNA may be under circadian control. mRNA transcripts may display different isoforms exhibiting varying 3'UTR lengths that have arisen by multiple polyadenylation sites in a process named alternative polyadenylation (Derti et al., 2012; Tian, Hu, Zhang, & Lutz, 2005). This alternative polyadenylation may affect mRNA stability, translation and subcellular localization, since the 3'UTR is a region that contains specific sequences recognized by microRNAs (miRNAs) and RBPs (Di Giammartino, Nishida, & Manley, 2011; Lai, 2002). Two proteins, cold-inducible RNA-binding protein (CIRBP) and RNA-binding motif protein 3 (RBM3), are known to be involved in the choice of the Poly(A) site (Liu et al., 2013b). Together, they mediate the

-WILEY WIREs 7 of 13

temperature synchronization of peripheral clocks in murine liver and embryonic fibroblasts (Morf et al., 2012). The expression of these two RBPs is thus temperature dependent. Under cold-shock conditions, CIRBP and RBM3 are up-regulated, leading to the preferential expression of transcripts with long 3'UTRs, which are more stable. By contrast, the depletion of CIRBP or RBM3 results in the production of preferential short 3'UTR variants (Liu et al., 2013b). A lot of transcripts in the liver have been shown to exhibit a circadian rhythm driven by the alternative use of proximal versus distal polyadenylation sites (Liu et al., 2013b).

#### 1.9 | RNA transport: nuclear retention and export

In addition to the different steps of RNA processing, RNA nucleocytoplasmic trafficking may also be under circadian timekeeping control (see Figure 2).

#### 1.9.1 | Nuclear retention

It was shown recently that the nuclear retention of mRNAs follows a circadian profile. RNAs can be rhythmically retained in the nucleus by particular bodies called paraspeckles (Torres et al., 2016; Torres et al., 2017). These nuclear bodies are formed around the long non-coding RNA, the nuclear enriched abundant transcript 1 (Neat1), and several RBPs. Paraspeckles are believed to control gene expression at the post-transcriptional level through the nuclear retention of mRNAs which contain in their 3'UTR inverted repeats of Alu sequences (IRAlu) (Chen & Carmichael, 2009; Chen, DeCerbo, & Carmichael, 2008). It was shown that the expression of all components of paraspeckles follows a circadian pattern that leads to rhythmic variations in paraspeckle number within the cells. Finally, thanks to their circadian expression pattern and given their functions in mRNA nuclear retention, it was shown that paraspeckles rhythmically retain RNAs in the nucleus. This rhythmic nuclear retention leads to the rhythmic expression of the corresponding genes (Torres et al., 2016; Torres et al., 2017) (see Figure 2).

#### 1.9.2 | RNA editing

It is known that mRNAs with 3'UTR IRAlu sequences can be A-to-I edited since IRAlu can give rise to double strand RNA (dsRNA). In animals, the most common editing event is adenosine to inosine (A-to-I) RNA editing, which converts

FIGURE 2 Circadian RNA transport and editing. In the nucleus mRNA may be rhythmically retained by paraspeckles. These nuclear bodies that are formed around a long non-coding RNA, Nuclear Enriched Abundant Transcript 1 (Neat1, purple lines), and several RNA-binding proteins (colored circles), vary in number within the nucleus during a circadian period. Editing of mRNA can also follow a circadian pattern driven by the circadian expression of ADAR, an A-to-I RNA editing enzyme. mRNA containing double strand RNA (dsRNA) in their 3' UTR can be A-to-I edited and edited RNAs are preferentially retained in the nucleus by paraspeckles. The circadian timekeeping system can also control mRNA nuclear export by means of some RNA-binding proteins including Cold Inducible RNA Binding Protein (CIRBP), Fragile X mental retardation 1 (FMR1), Fragile X related gene 1 and 2 (FXR1 and FXR2) and probably other proteins that act like a cargo by forming complexes with mRNA



adenosine to inosine in dsRNA regions through the action of adenosine deaminase on RNA (ADAR) enzymes. Hyperedited RNAs are preferentially retained in the nucleus by paraspeckles (Chen et al., 2008; Chen & Carmichael, 2009; Zhang & Carmichael, 2001). A-to-I RNA editing affects RNA abundance through various post-transcriptional mechanisms (Liscovitch, Bazak, Levanon, & Chechik, 2014). Indeed, when RNA editing occurs on coding sequences, it leads to protein diversification, but when it occurs on non-coding regions, it is believed to play a role in gene silencing via processes such as nuclear retention or RNA degradation (Liscovitch et al., 2014). It is interesting that A-to-I RNA editing has been recently identified as a key mechanism of post-transcriptional regulation in the circadian clockwork (Terajima et al., 2017). Through its circadian expression pattern, ADARB1, an A-to-I RNA editing enzyme, has been shown to play a pivotal role in generating cyclic editing and in driving mRNA expression rhythms through regulation of mRNA stability (Terajima et al., 2017). Indeed, in the mouse liver, a number of mRNA rhythms are post-transcriptionally generated by ADARB1 through the circadian rhythm of A-to-I editing in a variety of transcripts (Terajima et al., 2017) (see Fig. 2).

#### 1.9.3 | Nuclear export

The opposite process of nuclear retention of mRNAs, namely their nuclear export, is also time-controlled by the circadian clock. This circadian control of mRNA nuclear export involves some RBPs, such as CIRBP, Fragile X mental retardation 1 (*Fmr1*), and Fragile X related genes 1 and 2 (*Fxr1* and *Fxr2*) (Dockendorff et al., 2002; Kim, Bellini, & Ceman, 2009; Morf et al., 2012; Zhang et al., 2008) (see Figure 2). CIRBP is thought to protect bound target RNAs from degradation and by forming complexes with several gene transcripts to facilitate their nuclear export (Morf et al., 2012). As mentioned above, CIRBP has also been shown to regulate the choice of alternative polyadenylation sites and the splicing efficiency (Brown et al., 2013a).

The expression of CIRBP follows a circadian pattern driven by temperature cycles (Morf et al., 2012). As mentioned above, the molecular mechanisms underpinning temperature-mediated CIRBP expression were recently shown to involve, in addition to temperature-dependent mRNA instability, pre-mRNA splicing efficiency in NIH3T3 fibroblasts (Gotic et al., 2016; Goto et al., 2017). The role of CIRBP in circadian gene expression was initially discovered by loss-of-function experiments, which demonstrated that depletion of CIRBP in cultured NIH3T3 cells resulted in the broad loss of high-amplitude rhythms of circadian gene expression and a less robust circadian clock that was more prone to resetting (Morf et al., 2012). Among the identified CIRBP target mRNAs, several were CGs, including *Clock*, one of the core genes of the TTFL. The data suggest that the effects of CIRBP knockdown on circadian gene expression are mediated by the altered nuclear export of *Clock* mRNA (Morf et al., 2012). Core body temperature rhythms can then drive rhythms in CIRBP levels, which rhythmically regulate *Clock* mRNA cytoplasmic accumulation. This post-transcriptional process enhances the robustness of the circadian clock mechanism (Morf et al., 2012).

FMR1, FXR1 and FXR2 are nuclear RBPs that have been shown to mediate the nuclear export of mRNAs (Kim et al., 2009). The loss of expression of *fmr1* in flies, or both *Fxr1* and *Fxr2* in mice, induces a corresponding loss of circadian locomotor activity (Dockendorff et al., 2002; Kim et al., 2009; Zhang et al., 2008).

#### 1.10 | Circadian regulation by miRNAs

The identified miRNA participation in circadian clock mechanisms adds the possibility of additional layers of posttranscriptional regulation of clock gene expression (see Figure 3).

In animals, microRNAs are a class of noncoding, single-stranded RNA molecules containing 22 nucleotides. Numerous reports have unveiled the importance of miRNAs in the regulation of the circadian system (for reviews, see (Mehta & Cheng, 2013; Wang, Tian, Li, & Zheng, 2015)). Using mice carrying conditional alleles for the Dicer1 gene, Du et al. showed that miRNA regulation could impact up to 30% of the cycling transcripts by affecting the phase and the amplitude of rhythmic mRNA accumulation (Du, Arpat, De Matos, & Gatfield, 2014). Cheng et al. implicated miR-219 and miR-132 in the time-keeping mechanism. They showed that the brain-specific miR-219 which shows daily oscillations in the SCN and which was shown to target the CLOCK-BMAL1 complex, was implicated in period determination, whereas miR-132 was implicated in photic entrainment, acting as a feedback inhibitor to photic responsiveness (Cheng et al., 2007). The same group more recently showed that SCN neurons from miR-132/212-deficient mice have reduced dendritic spine density, along with altered methyl CpG-binding protein (MeCP2) rhythms (Mendoza-Viveros et al., 2017). Genes involved in chromatin remodeling such as Mecp2 but also those involved in translational control are direct targets of miR-132 in the mouse (Mendoza-Viveros et al., 2017; Alvarez-Saavedra et al., 2011) In the SCN, miR-132 modulated CCG expression through a MAPK-CREB-dependent mechanism. Microarrays and RNA-seq analysis provided evidence that a number of miRNAs follow a circadian expression pattern in different peripheral oscillators like the retina and mouse liver (Na et al., 2009; Vollmers et al., 2012; Wang et al., 2016; Xu, Witmer, Lumayag, Kovacs, & Valle, 2007).

Furthermore, some of these circadian miRNAs target CGs. In particular, cell-based studies have implicated miR-192, miR-194 and miR-29 in the regulation of the human *Per* gene family (Nagel, Clijsters, & Agami, 2009) and miR-24, miR-

WILEY WIREs 9 of 13



FIGURE 3 miRNA targeting the circadian system in mouse and human. Some microRNAs regulate core-clock genes or clock-controlled genes. Among them, some follow a circadian expression pattern

29a and miR-30a in the regulation of the mouse *Per* gene family (Chen, D'Alessandro, & Lee, 2013; Zhao et al., 2014) (see Figure 3). In humans, miR-107 controls the circadian rhythm of cells by binding to the *Clock* gene (Daimiel-Ruiz et al., 2015). In mice, miR-494, miR-142-3p and miR-27b-3p may function as post-transcriptional modulators of *Bmal1*, but only miR-494 follows a circadian expression pattern (Shende, Goldrick, Ramani, & Earnest, 2011; Zhang et al., 2016). More globally, circadian rhythms were dramatically shortened by 2 hours in Dicer-deficient (and thus miRNA-deficient) cells and mice (Chen et al., 2013). In mice, miR-17-5p was shown to be an important factor in controlling the circadian period. This control involves the stabilization of the circadian-clock period by interaction with *Clock* and *Npas2* and CRY1 may be an important out-put signaling molecule of *Clock* in mediating the period shortening elicited by miR-17-5p (Gao, Zhou, Yang, & Cao, 2016). The rhythmicity of *Cry1* mRNA translation is also controlled in the mouse by miR-185 (Lee et al., 2013).

Two recent studies explored the differential expression of miRNAs targeting CGs in colorectal cancer cells compared to non-tumor cells (Mazzoccoli et al., 2016; Wu, Fesler, & Ju, 2016). The results reported a modulation of expression of some of these miRNAs and an inverse correlation between the expression of miR-192 and miR-194 and that of *Per1* and *Per2* (Wu et al., 2016). In human colon cancer cell lines, an inverse correlation was also observed between miR-139-5p, a candidate tumor suppressor that regulates proliferation, migration and apoptosis, and TIMELESS (Mazzoccoli et al., 2016).

Different miRNAs were also shown to play a role in circadian functioning by targeting CCGs. For instance, two miR-NAs, (miR-96 and miR-182) are directed against the 3'UTR of adenylate cyclase and oscillate in anti-phase with this mRNA, suggesting that they play a role in the rhythmic pattern of this gene (Xu et al., 2007). In mouse osteoblasts, miR-433 helps to maintain circadian rhythms by regulating sensitivity to glucocorticoid receptor signaling (Smith, Dole, Franceschetti, Hrdlicka, & Delany, 2016). In the mouse liver, miR-122 is implicated in the up- and down-regulation of hundreds of mRNAs and contributes to the circadian expression of some of them, a number of which have been implicated in hepatic lipid metabolism (Gatfield et al., 2009). However, the hepatocyte-specific miR-122, which is rhythmically transcribed under the control of REVERB $\alpha$ , accumulates at nearly constant rates throughout the circadian period due to its long half-life exceeding 24 h (Gatfield et al., 2009). The fact that constant miR-122 levels could shape the circadian rhythm of a target led the authors to postulate that the constant repression of translation by a miRNA could increase the amplitude of cycling and convert a low-amplitude mRNA rhythm into a higher amplitude protein rhythm (Gatfield et al., 2009). Of interest, it has been shown that the circadian deadenylase nocturnin belongs to the miR-122 targets (Kojima, Gatfield, Esau, & Green, 2010). By regulating the expression of nocturnin, miR-122 could have an important impact on the clock output pathway involving the deadenylation by nocturnin and then the stability of its mRNA targets. It thus appears that miRNAs, either by impacting the core molecular clock or various clock output pathways, represent an additional and important component of the posttranscriptional mechanisms that can affect the functioning of circadian timekeeping (see Figure 3).

#### 2 | CONCLUSION

As described in this review, the mammalian circadian timekeeping system is now known to have the ability to control multiple steps of the RNA life cycle, from RNA processing to RNA cytoplasmic transport and degradation. Given that de novo

## 10 of 13 WILEY WIRES

transcription accounts for a small minority of cycling transcripts, it has now become evident that post-transcriptional mechanisms play a crucial role in regulating various aspects of the circadian clock function. Pioneer studies have provided evidence for a circadian regulation of RNA processing usually driven by RBPs that can target either splicing sites or the 5'UTR as well as 3'UTR regions controlling capping and RNA stability. In addition to this control of RNA processing, more recent studies have highlighted other steps in the RNA life cycle that can be under circadian regulation, such as A-to-I editing, rhythmic nuclear retention by paraspeckle nuclear bodies or degradation by miRNAs.

However, circadian clock regulation is immensely more complex than was once thought. This complexity not only relies on the diverse post-transcriptional, translational and posttranslational mechanisms described until now but also on the fact that while these processes have often been considered isolated steps, they can also be highly inter-connected. This is exemplified by the complex temperature-dependent regulation exerted by CIRBP, which has been shown to regulate circadian gene expression by controlling different post-transcriptional steps, such as alternative polyadenylation (Liu et al., 2013b), splicing efficiency (Gotic et al., 2016; Gotic & Schibler, 2017) and nuclear export (Morf et al., 2012). It remains to be determined whether several inter-connected processes can regulate a single RNA species. Conceivably, the circadian timekeeping system must maintain the correct period, phase and amplitude of the rhythms of thousands of proteins that generate the wide range of rhythmic biological processes. Thus, the circadian gene expression both within cells and in the central and peripheral tissues must be tightly regulated at several steps.

#### CONFLICT OF INTEREST

The authors have declared no conflicts of interest for this article.

#### ORCID

Denis Becquet b http://orcid.org/0000-0001-8232-7418 Jean-Louis Franc b http://orcid.org/0000-0002-2900-5468 Anne-Marie François-Bellan b http://orcid.org/0000-0002-3278-4642

#### REFERENCES

- Abe, M., Herzog, E. D., Yamazaki, S., Straume, M., Tei, H., Sakaki, Y., ... Block, G. D. (2002). Circadian rhythms in isolated brain regions. *The Journal of Neuroscience*, 22(1), 350–356.
- Albrecht, U. (2017). The circadian clock, metabolism and obesity. Obesity Reviews, 18(Suppl 1), 25-33.
- Alvarez-Saavedra, M., Antoun, G., Yanagiya, A., Oliva-Hernandez, R., Cornejo-Palma, D., Perez-Iratxeta, C., ... Cheng, H. Y. (2011). MiRNA-132 orchestrates chromatin remodeling and translational control of the circadian clock. *Human Molecular Genetics*, 20(4), 731–751.
- Avitabile, D., Genovese, L., Ponti, D., Ranieri, D., Raffa, S., Calogero, A., & Torrisi, M. R. (2014). Nucleolar localization and circadian regulation of per2s, a novel splicing variant of the period 2 gene. *Cellular and Molecular Life Sciences*, 71(13), 2547–2559.
- Bailey, S. M., Udoh, U. S., & Young, M. (2014). Circadian regulation of metabolism. The Journal of Endocrinology, 222(2), R75-R96.
- Bartok, O., Kyriacou, C. P., Levine, J., Sehgal, A., & Kadener, S. (2013). Adaptation of molecular circadian clockwork to environmental changes: A role for alternative splicing and mirnas. *Proceedings of the Biological Sciences*, 280(1765), 20130011.
- Bell-Pedersen, D., Cassone, V. M., Earnest, D. J., Golden, S. S., Hardin, P. E., Thomas, T. L., & Zoran, M. J. (2005). Circadian rhythms from multiple oscillators: Lessons from diverse organisms. *Nature Reviews Genetics*, 6(7), 544–556.
- Buhr, E. D., Yoo, S. H., & Takahashi, J. S. (2010). Temperature as a universal resetting cue for mammalian circadian oscillators. Science, 330(6002), 379-385.
- Cao, R., Gkogkas, C. G., De Zavalia, N., Blum, I. D., Yanagiya, A., Tsukumo, Y., ... Liu, A. C. (2015). Light-regulated translational control of circadian behavior by eif4e phosphorylation. *Nature Neuroscience*, 18(6), 855–862.
- Cao, R., Robinson, B., Xu, H., Gkogkas, C., Khoutorsky, A., Alain, T., ... Sonenberg, N. (2013). Translational control of entrainment and synchrony of the suprachiasmatic circadian clock by mtor/4E-BP1 signaling. *Neuron*, 79(4), 712–724.
- Chen, L. L., & Carmichael, G. G. (2009). Altered nuclear retention of mrnas containing inverted repeats in human embryonic stem cells: Functional role of a nuclear noncoding RNA. *Molecular Cell*, 35(4), 467–478.

Chen, L. L., DeCerbo, J. N., & Carmichael, G. G. (2008). Alu element-mediated gene silencing. The EMBO Journal, 27(12), 1694–1705.

- Chen, M., & Manley, J. L. (2009). Mechanisms of alternative splicing regulation: Insights from molecular and genomics approaches. *Nature Reviews. Molecular Cell Biology*, 10(11), 741–754.
- Chen, R., D'Alessandro, M., & Lee, C. (2013). MiRNAs are required for generating a time delay critical for the circadian oscillator. *Current Biology*, 23(20), 1959–1968.
- Cheng, H. Y., Papp, J. W., Varlamova, O., Dziema, H., Russell, B., Curfman, J. P., ... Obrietan, K. (2007). MicroRNA modulation of circadian-clock period and entrainment. *Neuron*, 54(5), 813–829.
- Cohen, S. E., & Golden, S. S. (2015). Circadian rhythms in cyanobacteria. Microbiology and Molecular Biology Reviews, 79(4), 373-385.
- Daimiel-Ruiz, L., Klett-Mingo, M., Konstantinidou, V., Micó, V., Aranda, J. F., García, B., ... Ordovás, J. M. (2015). Dietary lipids modulate the expression of mir-107, an mirna that regulates the circadian system. *Molecular Nutrition & Food Research*, 59(3), 552–565.
- Derti, A., Garrett-Engele, P., Macisaac, K. D., Stevens, R. C., Sriram, S., Chen, R., ... Babak, T. (2012). A quantitative atlas of polyadenylation in five mammals. Genome Research, 22(6), 1173–1183.
- Di Giammartino, D. C., Nishida, K., & Manley, J. L. (2011). Mechanisms and consequences of alternative polyadenylation. Molecular Cell, 43(6), 853-866.
- Dockendorff, T. C., Su, H. S., McBride, S. M., Yang, Z., Choi, C. H., Siwicki, K. K., ... Jongens, T. A. (2002). Drosophila lacking dfmr1 activity show defects in circadian output and fail to maintain courtship interest. *Neuron*, 34(6), 973–984.



Du, N. H., Arpat, A. B., De Matos, M., & Gatfield, D. (2014). MicroRNAs shape circadian hepatic gene expression on a transcriptome-wide scale. eLife, 3, e02510.

Duffield, G. E. (2003). DNA microarray analyses of circadian timing: The genomic basis of biological time. *Journal of Neuroendocrinology*, *15*(10), 991–1002. Duong, H. A., Robles, M. S., Knutti, D., & Weitz, C. J. (2011). A molecular mechanism for circadian clock negative feedback. *Science*, *332*(6036), 1436–1439.

Gao, Q., Zhou, L., Yang, S. Y., & Cao, J. M. (2016). A novel role of microrna 17-5p in the modulation of circadian rhythm. *Scientific Reports*, *6*, 30070.

- Garbarino-Pico, E., Niu, S., Rollag, M. D., Strayer, C. A., Besharse, J. C., & Green, C. B. (2007). Immediate early response of the circadian polya ribonuclease nocturnin to two extracellular stimuli. RNA, 13(5), 745–755.
- Gatfield, D., Le Martelot, G., Vejnar, C. E., Gerlach, D., Schaad, O., Fleury-Olela, F., ... Schibler, U. (2009). Integration of microrna mir-122 in hepatic circadian gene expression. Genes & Development, 23(11), 1313–1326.
- Gerstner, J. R., Vanderheyden, W. M., LaVaute, T., Westmark, C. J., Rouhana, L., Pack, A. I., ... Landry, C. F. (2012). Time of day regulates subcellular trafficking, tripartite synaptic localization, and polyadenylation of the astrocytic fabp7 mrna. *The Journal of Neuroscience*, *32*(4), 1383–1394.
- Gotic, I., & Schibler, U. (2017). Posttranscriptional mechanisms controlling diurnal gene expression cycles by body temperature rhythms. *RNA Biology*, 14(10), 1294–1298.
- Gotic, I., Omidi, S., Fleury-Olela, F., Molina, N., Naef, F., & Schibler, U. (2016). Temperature regulates splicing efficiency of the cold-inducible rna-binding protein gene cirbp. *Genes & Development*, 30(17), 2005–2017.
- Goto, M., Mizuno, M., Matsumoto, A., Yang, Z., Jimbo, E. F., Tabata, H., ... Nagata, K. I. (2017). Role of a circadian-relevant gene NR1D1 in brain development: Possible involvement in the pathophysiology of autism spectrum disorders. *Scientific Reports*, 7, 43945.
- Guilding, C., Hughes, A. T., Brown, T. M., Namvar, S., & Piggins, H. D. (2009). A riot of rhythms: Neuronal and glial circadian oscillators in the mediobasal hypothalamus. *Molecular Brain*, 2, 28.
- Guillaumond, F., Boyer, B., Becquet, D., Guillen, S., Kuhn, L., Garin, J., ... François-Bellan, A. M. (2011). Chromatin remodeling as a mechanism for circadian prolactin transcription: Rhythmic NONO and SFPQ recruitment to HLTF. *The FASEB Journal*, 25(8), 2740–2756.
- Guillaumond, F., Dardente, H., Giguère, V., & Cermakian, N. (2005). Differential control of bmall circadian transcription by REV-ERB and ROR nuclear receptors. *Journal of Biological Rhythms*, 20(5), 391–403.
- Harmer, S. L., Panda, S., & Kay, S. A. (2001). Molecular bases of circadian rhythms. Annual Review of Cell and Developmental Biology, 17(1), 215–253.
- Hastings, M. H., Maywood, E. S., & Reddy, A. B. (2008). Two decades of circadian time. Journal of Neuroendocrinology, 20(6), 812-819.
- Hernando, C. E., Romanowski, A., & Yanovsky, M. J. (2017). Transcriptional and post-transcriptional control of the plant circadian gene regulatory network. *Biochi*mica et Biophysica Acta, 1860(1), 84–94.
- Heyd, F., & Lynch, K. W. (2010). Phosphorylation-dependent regulation of PSF by GSK3 controls CD45 alternative splicing. Molecular Cell, 40(1), 126–137.
- Hong, C. I., Zámborszky, J., Baek, M., Labiscsak, L., Ju, K., Lee, H., ... Csikász-Nagy, A. (2014). Circadian rhythms synchronize mitosis in neurospora crassa. Proceedings of the National Academy of Sciences of the United States of America, 111(4), 1397–1402.
- Hughes, M. E., Grant, G. R., Paquin, C., Qian, J., & Nitabach, M. N. (2012). Deep sequencing the circadian and diurnal transcriptome of drosophila brain. *Genome Research*, 22(7), 1266–1281.
- Hurley, J., Loros, J. J., & Dunlap, J. C. (2015). Dissecting the mechanisms of the clock in neurospora. Methods in Enzymology, 551, 29-52.
- Hurley, J. M., Dasgupta, A., Emerson, J. M., Zhou, X., Ringelberg, C. S., Knabe, N., ... Dunlap, J. C. (2014). Analysis of clock-regulated genes in neurospora reveals widespread posttranscriptional control of metabolic potential. *Proceedings of the National Academy of Sciences of the United States of America*, 111(48), 16995–17002.
- Jagannath, A., Hughes, S., Abdelgany, A., Pothecary, C. A., Di Pretoro, S., Pires, S. S., ... Peirson, S. N. (2015). Isoforms of melanopsin mediate different behavioral responses to light. *Current Biology*, 25(18), 2430–2434.
- Jalkanen, A. L., Coleman, S. J., & Wilusz, J. (2014). Determinants and implications of mrna poly(a) tail size--does this protein make my tail look big? Seminars in Cell & Developmental Biology, 34, 24–32.
- Kadener, S., Menet, J. S., Sugino, K., Horwich, M. D., Weissbein, U., Nawathean, P., ... Rosbash, M. (2009). A role for micrornas in the drosophila circadian clock. Genes & Development, 23(18), 2179–2191.
- Kettner, N. M., Katchy, C. A., & Fu, L. (2014). Circadian gene variants in cancer. Annals of Medicine, 46(4), 208-220.
- Kim, D. Y., Kwak, E., Kim, S. H., Lee, K. H., Woo, K. C., & Kim, K. T. (2011). HnRNP Q mediates a phase-dependent translation-coupled mrna decay of mouse period3. Nucleic Acids Research, 39(20), 8901–8914.
- Kim, D. Y., Woo, K. C., Lee, K. H., Kim, T. D., & Kim, K. T. (2010). HnRNP Q and PTB modulate the circadian oscillation of mouse rev-erb alpha via ires-mediated translation. *Nucleic Acids Research*, 38(20), 7068–7078.
- Kim, H. J., Lee, H. R., Seo, J. Y., Ryu, H. G., Lee, K. H., Kim, D. Y., & Kim, K. T. (2017). Heterogeneous nuclear ribonucleoprotein A1 regulates rhythmic synthesis of mouse nfil3 protein via ires-mediated translation. *Scientific Reports*, 7, 42882.
- Kim, M., Bellini, M., & Ceman, S. (2009). Fragile X mental retardation protein FMRP binds mrnas in the nucleus. *Molecular and Cellular Biology*, 29(1), 214–228.
- Kim, S. H., Lee, K. H., Kim, D. Y., Kwak, E., Kim, S., & Kim, K. T. (2015). Rhythmic control of mrna stability modulates circadian amplitude of mouse period3 mrna. Journal of Neurochemistry, 132(6), 642–656.
- Kim, T. D., Woo, K. C., Cho, S., Ha, D. C., Jang, S. K., & Kim, K. T. (2007). Rhythmic control of AANAT translation by hnrnp Q in circadian melatonin production. Genes & Development, 21(7), 797–810.
- Koike, N., Yoo, S. H., Huang, H. C., Kumar, V., Lee, C., Kim, T. K., & Takahashi, J. S. (2012). Transcriptional architecture and chromatin landscape of the core circadian clock in mammals. *Science*, 338(6105), 349–354.
- Kojima, S., Gatfield, D., Esau, C. C., & Green, C. B. (2010). MicroRNA-122 modulates the rhythmic expression profile of the circadian deadenylase nocturnin in mouse liver. PLoS One, 5(6), e11264.
- Kojima, S., Matsumoto, K., Hirose, M., Shimada, M., Nagano, M., Shigeyoshi, Y., ... Tei, H. (2007). LARK activates posttranscriptional expression of an essential mammalian clock protein, PERIOD1. Proceedings of the National Academy of Sciences of the United States of America, 104(6), 1859–1864.
- Kojima, S., Sher-Chen, E. L., & Green, C. B. (2012). Circadian control of mrna polyadenylation dynamics regulates rhythmic protein expression. Genes & Development, 26(24), 2724–2736.
- Kowalska, E., Ripperger, J. A., Muheim, C., Maier, B., Kurihara, Y., Fox, A. H., ... Brown, S. A. (2012). Distinct roles of DBHS family members in the circadian transcriptional feedback loop. *Molecular and Cellular Biology*, 32(22), 4585–4594.
- Kwak, E., Kim, T. D., & Kim, K. T. (2006). Essential role of 3'-untranslated region-mediated mrna decay in circadian oscillations of mouse period3 mrna. The Journal of Biological Chemistry, 281(28), 19100–19106.
- Lai, E. C. (2002). Micro rnas are complementary to 3' UTR sequence motifs that mediate negative post-transcriptional regulation. Nature Genetics, 30(4), 363-364.
- Lee, K. -H., Kim, S. -H., Kim, D. -Y., Kim, S., & Kim, K. -T. (2012). Internal ribosomal entry site-mediated translation is important for rhythmic PERIOD1 expression. *PLoS One*, 7(5), e37936.
- Lee, K. H., Kim, S. H., Lee, H. R., Kim, W., Kim, D. Y., Shin, J. C., ... Kim, K. T. (2013). MicroRNA-185 oscillation controls circadian amplitude of mouse cryptochrome 1 via translational regulation. *Molecular Biology of the Cell*, 24(14), 2248–2255.

## 12 of 13 WILEY WIRES

- Lee, K. H., Woo, K. C., Kim, D. Y., Kim, T. D., Shin, J., Park, S. M., ... Kim, K. T. (2011). Rhythmic interaction between period1 mrna and hnrnp Q leads to circadian time-dependent translation. Biol: Mol. Cell.
- Lerner, I., Bartok, O., Wolfson, V., Menet, J. S., Weissbein, U., Afik, S., ... Kadener, S. (2015). Clk post-transcriptional control denoises circadian transcription both temporally and spatially. *Nature Communications*, *6*, 7056.
- Li, J., Grant, G. R., Hogenesch, J. B., & Hughes, M. E. (2015). Considerations for rna-seq analysis of circadian rhythms. Methods in Enzymology, 551, 349-367.
- Lipton, J. O., Yuan, E. D., Boyle, L. M., Ebrahimi-Fakhari, D., Kwiatkowski, E., Nathan, A., ... Sahin, M. (2015). The circadian protein BMAL1 regulates translation in response to s6k1-mediated phosphorylation. Cell, 161(5), 1138–1151.
- Liscovitch, N., Bazak, L., Levanon, E. Y., & Chechik, G. (2014). Positive correlation between ADAR expression and its targets suggests a complex regulation mediated by RNA editing in the human brain. RNA Biology, 11(11), 1447–1456.
- Liu, S., Brown, J. D., Stanya, K. J., Homan, E., Leidl, M., Inouye, K., ... Lee, C. H. (2013a). A diurnal serum lipid integrates hepatic lipogenesis and peripheral fatty acid use. *Nature*, 502(7472), 550–554.
- Liu, Y., Hu, W., Murakawa, Y., Yin, J., Wang, G., Landthaler, M., & Yan, J. (2013b). Cold-induced rna-binding proteins regulate circadian gene expression by controlling alternative polyadenylation. Scientific Reports, 3, 2054.
- Lowrey, P. L., & Takahashi, J. S. (2004). Mammalian circadian biology: Elucidating genome-wide levels of temporal organization. Annual Review of Genomics and Human Genetics, 5, 407–441.
- Mazzoccoli, G., Colangelo, T., Panza, A., Rubino, R., Tiberio, C., Palumbo, O., ... Mazza, T. (2016). Analysis of clock gene-mirna correlation networks reveals candidate drivers in colorectal cancer. *Oncotarget*, 7(29), 45444–45461.
- McGlincy, N. J., Valomon, A., Chesham, J. E., Maywood, E. S., Hastings, M. H., & Ule, J. (2012). Regulation of alternative splicing by the circadian clock and food related cues. *Genome Biology*, 13(6), R54.
- Mehta, N., & Cheng, H. Y. (2013). Micro-managing the circadian clock: The role of micromas in biological timekeeping. Journal of Molecular Biology, 425(19), 3609–3624.
- Menaker, M., Murphy, Z. C., & Sellix, M. T. (2013). Central control of peripheral circadian oscillators. Current Opinion in Neurobiology, 23(5), 741-746.
- Mendoza-Viveros, L., Chiang, C. K., Ong, J. L. K., Hegazi, S., Cheng, A. H., Bouchard-Cannon, P., ... Cheng, H. M. (2017). MiR-132/212 modulates seasonal adaptation and dendritic morphology of the central circadian clock. *Cell Reports*, 19(3), 505–520.
- Menet, J. S., Rodriguez, J., Abruzzi, K. C., & Rosbash, M. (2012). Nascent-Seq reveals novel features of mouse circadian transcriptional regulation. eLife, 1, e00011.
- Mezan, S., Ashwal-Fluss, R., Shenhav, R., Garber, M., & Kadener, S. (2013). Genome-wide assessment of post-transcriptional control in the fly brain. Frontiers in Molecular Neuroscience, 6(49), 1–9.
- Miller, B. H., McDearmon, E. L., Panda, S., Hayes, K. R., Zhang, J., Andrews, J. L., ... Takahashi, J. S. (2007). Circadian and clock-controlled regulation of the mouse transcriptome and cell proliferation. *Proceedings of the National Academy of Sciences of the United States of America*, 104(9), 3342–3347.
- Morf, J., Rey, G., Schneider, K., Stratmann, M., Fujita, J., Naef, F., & Schibler, U. (2012). Cold-inducible rna-binding protein modulates circadian gene expression posttranscriptionally. Science, 338(6105), 379–383.
- Morse, D., Milos, P. M., Roux, E., & Hastings, J. W. (1989). Circadian regulation of bioluminescence in gonyaulax involves translational control. *Proceedings of the National Academy of Sciences*, 86(1), 172–176.
- Musiek, E. S. (2015). Circadian clock disruption in neurodegenerative diseases: Cause and effect? Frontiers in Pharmacology, 6(29), 1-6.
- Na, Y. J., Sung, J. H., Lee, S. C., Lee, Y. J., Choi, Y. J., Park, W. Y., ... Kim, J. H. (2009). Comprehensive analysis of microrna-mrna co-expression in circadian rhythm. *Experimental & Molecular Medicine*, 41(9), 638–647.
- Nagel, R., Clijsters, L., & Agami, R. (2009). The mirna-192/194 cluster regulates the period gene family and the circadian clock. *The FEBS Journal*, 276(19), 5447–5455.
- Novák, B., & Tyson, J. J. (2008). Design principles of biochemical oscillators. Nature Reviews. Molecular Cell Biology, 9(12), 981–991.
- Ozkaya, O., & Rosato, E. (2012). The circadian clock of the fly: A neurogenetics journey through time. *Advances in Genetics*, 77, 79–123.
- Panda, S., Antoch, M. P., Miller, B. H., Su, A. I., Schook, A. B., Straume, M., ... Hogenesch, J. B. (2002). Coordinated transcription of key pathways in the mouse by the circadian clock. *Cell*, 109(3), 307–320.
- Pembroke, W. G., Babbs, A., Davies, K., Ponting, C. P., & Oliver, P. L. (2015). Temporal transcriptomics suggest that twin-peaking genes reset the clock. *eLife*, 4, e10518.
- Pires, S. S., Hughes, S., Turton, M., Melyan, Z., Peirson, S. N., Zheng, L., ... Halford, S. (2009). Differential expression of two distinct functional isoforms of melanopsin (opn4) in the mammalian retina. *The Journal of Neuroscience*, 29(39), 12332–12342.
- Potter, G. D., Cade, J. E., Grant, P. J., & Hardie, L. J. (2016). Nutrition and the circadian system. The British Journal of Nutrition, 116(3), 434-442.
- Preitner, N., Damiola, F., Lopez-Molina, L., Zakany, J., Duboule, D., Albrecht, U., & Schibler, U. (2002). The orphan nuclear receptor rev-erbalpha controls circadian transcription within the positive limb of the mammalian circadian oscillator. *Cell*, 110(2), 251–260.
- Preußner, M., Goldammer, G., Neumann, A., Haltenhof, T., Rautenstrauch, P., Müller-McNicoll, M., & Heyd, F. (2017). Body temperature cycles control rhythmic alternative splicing in mammals. *Molecular Cell*, 67(3), 433–446.e4.
- Preußner, M., Wilhelmi, I., Schultz, A. S., Finkernagel, F., Michel, M., Möröy, T., & Heyd, F. (2014). Rhythmic u2af26 alternative splicing controls PERIOD1 stability and the circadian clock in mice. *Molecular Cell*, 54(4), 651–662.
- Reddy, A. B., Karp, N. A., Maywood, E. S., Sage, E. A., Deery, M., O'Neill, J. S., ... Hastings, M. H. (2006). Circadian orchestration of the hepatic proteome. Current Biology, 16(11), 1107–1115.
- Robinson, B. G., Frim, D. M., Schwartz, W. J., & Majzoub, J. A. (1988). Vasopressin mrna in the suprachiasmatic nuclei: Daily regulation of polyadenylate tail length. *Science*, 241(4863), 342–344.
- Rodriguez, J., Tang, C. H., Khodor, Y. L., Vodala, S., Menet, J. S., & Rosbash, M. (2013). Nascent-Seq analysis of drosophila cycling gene expression. Proceedings of the National Academy of Sciences of the United States of America, 110(4), E275–E284.
- Saraf, A., Luo, J., Morris, D. R., & Storm, D. R. (2014). Phosphorylation of eukaryotic translation initiation factor 4E and eukaryotic translation initiation factor 4e-binding protein (4EBP) and their upstream signaling components undergo diurnal oscillation in the mouse hippocampus: Implications for memory persistence. *The Journal of Biological Chemistry*, 289(29), 20129–20138.
- Sato, T. K., Panda, S., Miraglia, L. J., Reyes, T. M., Rudic, R. D., McNamara, P., ... Hogenesch, J. B. (2004). A functional genomics strategy reveals rora as a component of the mammalian circadian clock. *Neuron*, 43(4), 527–537.
- Schibler, U., & Sassone-Corsi, P. (2002). A web of circadian pacemakers. Cell, 111(7), 919-922.
- Sellix, M. T. (2015). Circadian clock function in the mammalian ovary. Journal of Biological Rhythms, 30(1), 7–19.
- Shende, V. R., Goldrick, M. M., Ramani, S., & Earnest, D. J. (2011). Expression and rhythmic modulation of circulating micrornas targeting the clock gene bmall in mice. *PLoS One*, *6*(7), e22586.
- Smith, S. S., Dole, N. S., Franceschetti, T., Hrdlicka, H. C., & Delany, A. M. (2016). MicroRNA-433 dampens glucocorticoid receptor signaling, impacting circadian rhythm and osteoblastic gene expression. *The Journal of Biological Chemistry*, 291(41), 21717–21728.
- Soták, M., Sumová, A., & Pácha, J. (2014). Cross-talk between the circadian clock and the cell cycle in cancer. Annals of Medicine, 46(4), 221–232.

- Storch, K. -F., Lipan, O., Leykin, I., Viswanathan, N., Davis, F. C., Wong, W. H., & Weitz, C. J. (2002). Extensive and divergent circadian gene expression in liver and heart. *Nature*, 417(6884), 78–83.
- Takahashi, J. S. (2015). Molecular components of the circadian clock in mammals. Diabetes, Obesity & Metabolism, 17(Suppl 1), 6-11.
- Takeda, N., & Maemura, K. (2015). The role of clock genes and circadian rhythm in the development of cardiovascular diseases. Cellular and Molecular Life Sciences, 72(17), 3225–3234.
- Terajima, H., Yoshitane, H., Ozaki, H., Suzuki, Y., Shimba, S., Kuroda, S., ... Fukada, Y. (2017). ADARB1 catalyzes circadian a-to-i editing and regulates RNA rhythm. *Nature Genetics*, 49(1), 146–151.
- Tian, B., Hu, J., Zhang, H., & Lutz, C. S. (2005). A large-scale analysis of mrna polyadenylation of human and mouse genes. *Nucleic Acids Research*, 33(1), 201–212.
- Tonsfeldt, K. J., & Chappell, P. E. (2012). Clocks on top: The role of the circadian clock in the hypothalamic and pituitary regulation of endocrine physiology. *Molecular and Cellular Endocrinology*, 349(1), 3–12.
- Torres, M., Becquet, D., Blanchard, M. P., Guillen, S., Boyer, B., Moreno, M., ... François-Bellan, A. M. (2016). Circadian RNA expression elicited by 3'-UTR iralu-paraspeckle associated elements. *eLife*, 5, e14837.
- Torres, M., Becquet, D., Blanchard, M. P., Guillen, S., Boyer, B., Moreno, M., ... François-Bellan, A. M. (2017). Paraspeckles as rhythmic nuclear mrna anchorages responsible for circadian gene expression. *Nucleus*, 8(3), 249–254.
- Ueda, H. R., Hayashi, S., Chen, W., Sano, M., Machida, M., Shigeyoshi, Y., ... Hashimoto, S. (2005). System-level identification of transcriptional circuits underlying mammalian circadian clocks. *Nature Genetics*, 37(2), 187–192.
- Videnovic, A., Lazar, A. S., Barker, R. A., & Overeem, S. (2014). 'The clocks that time us'--circadian rhythms in neurodegenerative disorders. Nature Reviews. Neurology, 10(12), 683–693.
- Vollmers, C., Gill, S., DiTacchio, L., Pulivarthy, S. R., Le, H. D., & Panda, S. (2009). Time of feeding and the intrinsic circadian clock drive rhythms in hepatic gene expression. Proceedings of the National Academy of Sciences, 106(50), 21453–21458.
- Vollmers, C., Schmitz, R. J., Nathanson, J., Yeo, G., Ecker, J. R., & Panda, S. (2012). Circadian oscillations of protein-coding and regulatory mas in a highly dynamic mammalian liver epigenome. *Cell Metabolism*, 16(6), 833–845.
- Wang, H., Fan, Z., Zhao, M., Li, J., Lu, M., Liu, W., ... Yan, J. (2016). Oscillating primary transcripts harbor mirnas with circadian functions. Scientific Reports, 6, 21598.
- Wang, X., Tian, G., Li, Z., & Zheng, L. (2015). The crosstalk between mirna and mammalian circadian clock. Current Medicinal Chemistry, 22(13), 1582–1588.
- Wang, Y., Osterbur, D. L., Megaw, P. L., Tosini, G., Fukuhara, C., Green, C. B., & Besharse, J. C. (2001). Rhythmic expression of nocturnin mrna in multiple tissues of the mouse. BMC Developmental Biology, 1, 9.
- Weill, L., Belloc, E., Bava, F. A., & Méndez, R. (2012). Translational control by changes in poly(a) tail length: Recycling mrnas. Nature Structural & Molecular Biology, 19(6), 577–585.
- Woo, K. C., Ha, D. C., Lee, K. H., Kim, D. Y., Kim, T. D., & Kim, K. T. (2010). Circadian amplitude of cryptochrome 1 is modulated by mrna stability regulation via cytoplasmic hnrnp D oscillation. *Molecular and Cellular Biology*, 30(1), 197–205.
- Woo, K. C., Kim, T. D., Lee, K. H., Kim, D. Y., Kim, W., Lee, K. Y., & Kim, K. T. (2009). Mouse period 2 mrna circadian oscillation is modulated by ptb-mediated rhythmic mrna degradation. *Nucleic Acids Research*, 37(1), 26–37.
- Wu, S., Fesler, A., & Ju, J. (2016). Implications of circadian rhythm regulation by micrornas in colorectal cancer. Cancer Transl Med, 2(1), 1-6.
- Xu, S., Witmer, P. D., Lumayag, S., Kovacs, B., & Valle, D. (2007). MicroRNA (mirna) transcriptome of mouse retina and identification of a sensory organ-specific mirna cluster. *The Journal of Biological Chemistry*, 282(34), 25053–25066.
- Yoo, S. H., Yamazaki, S., Lowrey, P. L., Shimomura, K., Ko, C. H., Buhr, E. D., ... Takahashi, J. S. (2004). PERIOD2::LUCIFERASE real-time reporting of circadian dynamics reveals persistent circadian oscillations in mouse peripheral tissues. Proceedings of the National Academy of Sciences of the United States of America, 101(15), 5339–5346.
- Zhang, J., Fang, Z., Jud, C., Vansteensel, M. J., Kaasik, K., Lee, C. C., ... Nelson, D. L. (2008). Fragile x-related proteins regulate mammalian circadian behavioral rhythms. American Journal of Human Genetics, 83(1), 43–52.
- Zhang, R., Lahens, N. F., Ballance, H. I., Hughes, M. E., & Hogenesch, J. B. (2014). A circadian gene expression atlas in mammals: Implications for biology and medicine. Proceedings of the National Academy of Sciences of the United States of America, 111(45), 16219–16224.
- Zhang, W., Wang, P., Chen, S., Zhang, Z., Liang, T., & Liu, C. (2016). Rhythmic expression of mir-27b-3p targets the clock gene bmal1 at the posttranscriptional level in the mouse liver. *The FASEB Journal*, 30(6), 2151–2160.
- Zhang, Z., & Carmichael, G. G. (2001). The fate of dsrna in the nucleus: A p54(nrb)-containing complex mediates the nuclear retention of promiscuously a-to-i edited rnas. Cell, 106(4), 465–475.
- Zhao, X., Zhu, X., Cheng, S., Xie, Y., Wang, Z., Liu, Y., ... Wang, Y. (2014). MiR-29a/b/c regulate human circadian gene hper1 expression by targeting its 3'UTR. Acta Biochim Biophys Sin (Shanghai), 46(4), 313–317.
- Zwiebel, L. J., Hardin, P. E., Liu, X., Hall, J. C., & Rosbash, M. (1991). A post-transcriptional mechanism contributes to circadian cycling of a per-beta-galactosidase fusion protein. Proceedings of the National Academy of Sciences of the United States of America, 88(9), 3882–3886.

How to cite this article: Torres M, Becquet D, Franc J-L, François-Bellan A-M. Circadian processes in the RNA life cycle. *WIREs RNA*. 2018;e1467. <u>https://doi.org/10.1002/wrna.1467</u>

## II - Le long ARN non codant Neat1 et les paraspeckles

## 1. Les longs ARN non codants

## 1.1. Définition générale des longs ARN non codants

Le génome contient l'intégralité des plans nécessaires au fonctionnement des êtres vivants, leur développement et leur pérennisation. Dans le but de comprendre le fonctionnement et la composition du génome humain, un projet à échelle internationale a vu le jour. Ce projet appelé ENCODE, pour « Encyclopedia of DNA Elements », avait pour but l'étude de l'organisation des 3 milliards de bases qui composent le génome humain (ENCODE Project Consortium 2012). A l'issue de ce projet titanesque, il est apparu que les 3,2 milliards de bases qui composent le génome humain ne comprennent que 20 000 gènes codants pour des protéines. Le vaste nombre de bases restantes a longtemps été décrit comme de l'ADN *poubelle* composé de transposons et de pseudogènes. L'analyse du projet ENCODE a majoritairement contribué à démontrer qu'une large proportion de cet ADN est transcrit en ARNs. Ces ARNs ne codant pas pour des protéines ont été appelés ARN non codants.

Au sein des ARN non codants se distinguent deux catégories, les non-régulants et les régulants. On trouve parmi les ARN non-régulants, les ARN de transfert (ARNt) et les ARN ribosomiques (ARNr) connus pour leur rôle dans la synthèse des protéines, ainsi que les petits ARN nucléaires (snARN) et nucléolaires (snoARN). La classe des ARN régulants comprend les siARN, miARN et piARN ainsi que les longs ARN non codants (lncRNA).

Grâce à l'étude des données issues du projet ENCODE, le nombre de ces lncRNAs est évalué à 14 000 chez l'humain. Ce nombre élevé suggère un rôle physiologique supérieur à celui que les quelques exemples de lncRNAs étudiés laissaient imaginer.

Deux critères définissent l'appartenance d'un ARN à la classe des lncRNAs, sa longueur qui doit être supérieure à 200 nucléotides et l'absence de cadre de lecture (ORF). Ces critères peu sélectifs font que les lncRNAs comprennent une grande diversité d'ARNs tant au niveau de leurs fonctions que de leurs structures. Ainsi la taille des lncRNAs peut varier entre 200 nucléotides pour les plus petits et 100 000 nucléotides environ pour les plus longs (Redrup et al 2009).

Les lncRNAs peuvent aussi être sous-classés en fonction de la localisation de leur séquence dans le génome. Selon l'orientation de sa séquence le lncRNA sera sens ou anti-sens, et selon son positionnement par rapport à des séquences de gènes codants pour des protéines, il sera dit inter-génique si sa séquence est entre deux séquences codantes pour des protéines, ou génique si sa séquence chevauche une séquence codante (Derrien et al 2012).

Au niveau de leur expression, les lncRNAs semblent utiliser les mêmes voies que les ARNs codants pour des protéines. Ils ont les mêmes profils de modification par les histones, les mêmes codes de signalisation pour l'épissage et les mêmes longueurs moyennes d'intron et d'exons. Ils présentent néanmoins une différence dans leur nombre d'exons, entre 1 et 4 exons, contre 1 à 14 exons pour les ARNs codants pour des protéines (Derrien et al 2012). Le nombre d'isoformes pour un même gène est aussi inférieur à celui des gènes codants pour des protéines, avec 25% des gènes codants pour les lncRNAs ne produisant qu'une seule isoforme (Derrien et al 2012). Enfin l'analyse de l'expression des lncRNAs dans plusieurs organes humains montre qu'ils sont globalement moins exprimés que les gènes codants pour des protéines , que leur expression est tissu-spécifique avec un enrichissement plus fort dans le cerveau humain et qu'ils sont trouvés majoritairement dans le noyau des cellules (Derrien et al 2012).

Du point de vue évolutif, plusieurs études s'accordent sur l'existence d'une corrélation entre la complexité d'un organisme et la quantité de lncRNAs exprimés (Liu et al 2013; Kapusta, & Feschotte 2014; Derrien et al 2012) . L'apparition et l'augmentation de la quantité de lncRNAs seraient de ce fait un phénomène évolutif relativement récent. Par exemple, la comparaison du génome de 18 mammifères montre que seulement un tiers des lncRNAs présents chez l'humain est retrouvé chez les primates (Derrien et al 2012). Cette génération rapide des lncRNAs pourrait provenir de la forte redondance dans les mécanismes impliquant les lncRNAs, ce qui réduit grandement la pression sélective (Kapusta, & Feschotte 2014). Cette hypothèse est cohérente avec la plupart des études qui montrent des rôles variés des lncRNAs mais rarement l'apparition de problèmes physiologiques majeurs en cas d'inhibition de l'expression de l'un d'entre eux.

## 1.2. Les fonctions des longs ARN non codants

Le grand nombre de lncRNAs exprimés dans le génome humain suggère que cette catégorie d'ARNs non codants joue un rôle majeur dans le fonctionnement cellulaire. Cependant, l'étude des lncRNAs étant un domaine de recherche encore récent, peu de données existent concernant leurs fonctions, leurs modes d'action et leur importance dans les processus cellulaires. En effet, à ce jour seul un petit nombre de lncRNAs a été étudié en détail. A partir de ces études, il est apparu que les lncRNAs pourraient agir à différents niveaux de contrôle de l'expression des gènes.

Au niveau épigénétique, il a été montré que le lncRNA Xist participait à l'inhibition d'un des chromosomes X chez la femelle diploïde (Brown et al 1992). Le lncRNA Xist recouvre le chromosome X inactivé en formant une structure appelée Barr Body (Brown et al 1991). Il permet ainsi l'accumulation de plusieurs histones qui vont augmenter l'inactivation du chromosome X (Kalantry, & Magnuson 2006). Un autre lncRNA, HOTAIR (HOX transcript antisense RNA) agit aussi au niveau épigénétique pour réguler le locus HOXC (Rinn et al 2007). HOTAIR permet de recruter par ses extrémités 5' et 3' des protéines qui vont réprimer les modifications des histones et ainsi avoir un effet inhibant sur la transcription du locus HOXC (Tsai et al 2010).

Au niveau du contrôle de la transcription, il a été montré que le lncRNA LincRNA p21 recrute la protéine hnRNP-K pour former un complexe qui inhibe la transcription du gène cdkn1a codant pour la protéine p21 (Huarte et al 2010; Dimitrova et al 2014). Cette régulation participe à un mécanisme d'augmentation de l'apoptose dans les cellules.

Au niveau post-transcriptionnel, les lncRNAs peuvent agir sur différentes étapes de la maturation des ARNs. Il a ainsi été décrit que le lncRNA BACE1-AS joue un rôle sur la stabilité de l'ARNm BACE1. En effet le lncRNA BACE1-AS est la séquence antisens complémentaire de l'ARNm BACE1. Cette complémentarité permet la formation d'un complexe d'ARN double brin qui inhibe la fixation des complexes inhibiteurs des ARNs RISC, permettant alors d'augmenter la stabilité de l'ARNm BACE1 (Faghihi et al 2008). A un autre niveau du contrôle de l'expression des ARNm, il a été montré que le lncRNA LincRNA-p21 pouvait réguler la traduction en protéines de certains ARNm.



## Figure 6: Interactions et fonctions connues des longs ARN non codants (IncRNA)

Les lncRNAs sont associés à des protéines partenaires avec lesquelles ils peuvent intéragir avec l'ADN, les ARNs codants ou non codants, et de nombreuses autres protéines pour former des organelles. Au travers de ces intéractions, les lncRNAs sont capables de réguler l'expression des gènes à différents niveaux. Dans le noyau, LincRNA-p21 agit au niveau de transcription; en revanche dans le cytoplasme, il a été impliqué dans l'inhibition de la traduction de certains ARNm. Quand l'expression de LincRNA-p21 dans le cytoplasme est élevée, LincRNA-p21 s'associe à des polysomes et favorise l'association des ARNm avec des facteurs répresseurs de la traduction (Yoon et al 2012).

Enfin, les lncRNAs peuvent être intégrés à des structures nucléaires pour réguler l'expression des ARNs. C'est le cas du lncRNA Malat1 qui est localisé dans les speckles nucléaires (Hutchinson et al 2007). Malat1 est retenu par les protéines contenues dans les speckles nucléaires et permet de retenir à son tour d'autres protéines (Miyagawa et al 2012). Les speckles concentrent ainsi un grand nombre de protéines impliquées dans l'épissage des ARNs (Spector, & Lamond 2011). Malat1 en association avec ces protéines va contrôler l'épissage alternatif de nombreux ARNs. C'est l'interaction de Malat1 avec les protéines des speckles qui entraine leur activation par phosphorylation. Ces structures sont trouvées fréquemment à proximité de sites de transcription afin de permettre un épissage co-transcriptionnel.

Dans l'ensemble, ces études de cas illustrent les différents niveaux d'action des lncRNAs dans la régulation de l'expression des gènes (Fig6). Certains exemples, comme celui du lncRNA LincRNA-p21 qui agit au niveau transcriptionnel dans le noyau et post-transcriptionnel dans le cytoplasme, montrent qu'un même lncRNA peut agir à plusieurs niveaux. Le grand nombre de lncRNAs ainsi que leur évidente complexité de fonctionnement laisse entrevoir un rôle majeur dans le fonctionnement des organismes vivants qui les expriment.

## 2. Les paraspeckles

### 2.1. Découverte des paraspeckles

Les paraspeckles ont été observés pour la première fois en 1993 par microscopie électronique dans le noyau de cellules Hela. Ils sont alors décrits comme des granules denses localisés dans l'espace inter-chromatinien (IGAZ pour Interchromatin Granule Associated Zones (Visa et al 1993)) qui définissent un nouveau domaine nucléaire. Ce n'est qu'en 2002 qu'une étude des protéines présentes dans le noyau des cellules Hela permet d'identifier une protéine s'accumulant dans ce même domaine (Fox et al 2002) (fig7.A). Du fait de sa localisation à proximité d'un autre domaine inter-chromatinien appelé speckle, le nouveau domaine est appelé paraspeckle et la protéine ayant permis sa découverte est nommée paraspeckle component 1 (PSPC1 , anciennement appelé PSP1). La même équipe démontre la présence dans les paraspeckles de deux autres protéines, RNA binding motif protein 14 (RBM14, anciennement appelée PSP2) et Non-POU Domain Containing Octamer Binding (NONO, anciennement appelée p54nrb) (Bond, & Fox 2009).

En 2009 trois études montrent la présence du lncRNA Neat1 dans les paraspeckles (Clemson et al 2009; Sasaki et al 2009; Sunwoo et al 2009) . Il y est trouvé associé à PSPC1, NONO et une autre protéine Splicing Factor Proline and Glutamine-rich (SFPQ, anciennement appelée PSF) (Fig7.C). Neat1 existe sous deux isoformes, Neat1-1 et Neat1-2, toutes deux présentes dans les paraspeckles (Sasaki et al 2009). Les trois études démontrent que non seulement Neat1 est présent dans les paraspeckles, mais que son expression est indispensable à leur formation. En effet l'inhibition de l'expression de Neat1 induit inévitablement la disparition des paraspeckles.

## 2.2. Les éléments du paraspeckles

### 2.2.1. Le long ARN non codant Neat1

Le lncRNA Neat1 est l'un des lncRNAs les plus exprimés dans les cellules des mammifères (Gibb et al 2011). Il est codé par le gène Neat1 présent sur le chromosome 11q13 chez l'humain (Hutchinson et al 2007). L'analyse de la



D'après Sasaki et al., 2009



## Figure 7 : Localisation, composants principaux et structure des paraspeckles

A: Visualisation des paraspeckles dans des cellules HeLa YFP-PSPC1. Dans les noyaux des cellules marqués au DAPI (bleu), les paraspeckles sont localisés grâce à la fluorescence de la protéine PSPC1 (vert). Ils sont situés dans l'espace interchromatinien à proximité directe des speckles nucléaires marqués par immunofluorescence de la protéine SC35 (rouge).

B: Structure du gène Neat1. Chez l'homme le gène codant pour le lncRNA Neat1 est composé d'un seul exon qui est transcrit sous deux formes, une isoforme courte Neat1-1 et une isoforme longue Neat1-2. La polymérase est capable de passer au delà du signal de polyadénylation (Poly(A)) qui détermine la fin du premier transcrit, Neat1-1 (3,2 kb chez la souris, 3,7 kb chez l'homme), pour poursuivre dans la synthèse du deuxième, Neat1-2 (20 kb chez la souris, 23kb chez l'homme).

C: Colocalisation dans le noyau de cellules HeLa (marquage DAPI, bleu) des différents composants des paraspeckles. Neat1 (Men $\varepsilon/\beta$ ) est marqué par hybridation in situ en fluorescence et colocalise avec les protéines des paraspeckles NONO (p54), SFPQ (PSF) et PSPC1 (PSP1), marquées par immunofluorescence, pour former les focis caractéristiques des paraspeckles.

D: Modèle de structure des paraspeckles. Dans les paraspeckles, les isoformes Neat1-2 sont repliées avec leurs extrémités 5' et 3' regroupées séparemment et localisées dans la couche externe des paraspeckles (Shell). Les parties centrales des isoformes Neat1-2 sont concentrées dans la couche interne des paraspeckles (Core). Les différentes protéines des paraspeckles ont une distribution spécifique entre la couche externe, interne et la plaque intermédiaire (Patch).

conservation de la séquence du gène codant pour Neat1 au cours de l'évolution suggère que Neat1 pourrait provenir de la duplication du gène codant pour un autre lncRNA, Malat1, avec lequel il présente un fort alignement de séquence (Stadler 2010). Sachant que Malat1 est exprimé chez tous les vertébrés alors que Neat1 n'est présent que chez les mammifères, il est probable que Neat1 soit un gène apparu tardivement au cours de l'évolution.

Le promoteur de Neat1 qui recrute l'ARN polymérase II donne lieu à l'expression de deux isoformes (Hutchinson et al 2007; Sunwoo et al 2009; Sasaki et al 2009) . Ces deux isoformes sont de longueurs différentes avec une forme courte, Neat1-1 anciennement Menɛ (3,2kb chez la souris, 3.7kb chez l'humain), et une forme longue, Neat1-2 anciennement appelé Menß (20kb chez la souris, 23kb chez l'humain) (Sasaki et al 2009) (Fig7.B). Les isoformes ont une séquence commune en 5' et varient dans leur séquence en 3'. L'isoforme courte Neat1-1 est polyadénylée en 3'. Neat1-2 possède en 3' une structure similaire à un ARNt qui est reconnue et clivée par la RNAse P générant ainsi un transcrit qui n'est pas polyadénylé en 3' (Sunwoo et al 2009). L'isoforme Neat1-2 forme en 3' une triple hélice impliquée dans la stabilisation de cette isoforme (Brown et al 2012)

La production des deux isoformes de Neat1 est permise par les signaux de polyadénylation présents à la fin de la séquence codant pour l'isoforme Neat1-1. On sait maintenant que ces signaux sont reconnus par les protéines CPSF6 (cleavage et polyadenylation specific factor 6) et NUDT21 (Nudix type motif 21). Ces protéines se lient à la fin de la séquence codant pour Neat1-1 et induisent le clivage et la polyadénylation de cette isoforme, empêchant ainsi la transcription de l'isoforme Neat1-2 (Naganuma et al 2012). Pour que l'isoforme Neat1-2 soit transcrite il faut qu'il y ait un échec du clivage de Neat1-1. Il a ainsi été observé que l'utilisation de siRNA dirigés contre CPSF6 induit une augmentation de l'expression de l'isoforme Neat1-2. Naganuma et al. ont montré que dans la cellule, l'inhibition de CPSF6 pouvait provenir de la protéine HNRNPK (Heterogeneous nuclear ribonucleoprotein K). HNRNPK entrerait en compétition avec CPSF6 perturbant ainsi son action de clivage et permettrait la formation de Neat1-2 (Naganuma et al 2012).

Le ratio entre la transcription de l'isoforme courte Neat1-1 et l'isoforme longue Neat1-2 varie en fonction des tissus mais dans tous les tissus étudiés, l'isoforme Neat1-1 est exprimée au moins quatre fois plus que l'isoforme Neat1-2 (Sunwoo et al 2009; Nakagawa et al 2011; ENCODE Project Consortium 2012).

Neat1 est globalement exprimé dans la majorité des tissus chez la souris adulte (Nakagawa et al 2011). Les tissus embryonnaires en revanche ne présentent pas d'expression de Neat1, ce qui avait déjà été observé in vitro dans des lignées de cellules embryonnaires (Chen, & Carmichael 2009). Il est intéressant de noter que l'isoforme Neat1-1 a été observée en grande quantité dans tous les tissus alors que l'isoforme Neat1-2 présente des taux élevés d'expression particulièrement dans les tissus spécialisés dans les fonctions sécrétrices, comme les glandes salivaires, les surrénales et le tissu épithélial de l'estomac (Nakagawa et al 2011). Cette observation suggère des fonctions différentes pour les deux isoformes de Neat1.

### 2.2.2. Les protéines

Les protéines PSPC1, SFPQ et NONO ayant conduit à la découverte des paraspeckles appartiennent à la famille des protéines DBHS (Drosophila behavior human splicing) (Fox et al 2002). Elles possèdent des motifs de reconnaissance de l'ARN qui leur permettent de se fixer au lncRNA Neat1. Les protéines DBHS peuvent se lier à l'ADN double brin ainsi qu'à l'ARN simple et double brin. Elles vont ainsi pouvoir agir à de nombreux niveaux de contrôle de l'expression des gènes : initiation, co-activation, co-repression ou terminaison de la transcription de gènes, épissage et épissage alternatif des ARNs. Il a aussi été décrit que les DBHS pouvaient participer à la réparation de doubles brins d'ADN endommagés (Knott et al 2016).

Cependant ces protéines ne sont pas les seules à être présentes dans les paraspeckles. L'étude de Naganuma montre que dans la lignée cellulaire HeLa, sur plus de 18000 protéines étudiées, 35 protéines au total sont présentes dans les paraspeckles (Naganuma et al 2012). Cependant toutes ces protéines ne participent pas à la formation des paraspeckles. L'inhibition de l'expression soit de la protéine NONO, soit de la protéine SFPQ, induit une perte de la formation des paraspeckles (Sasaki et al 2009; Naganuma et al 2012). Il en va de même pour la protéine FUS (fused in sarcoma) (Naganuma et al 2012; Shelkovnikova et al 2014), ainsi que pour les protéines RBM14, HNRNPK (Heterogeneous nuclear ribonucleoprotein K), HNRNPH3 (Heterogeneous nuclear ribonucleoprotein H3) et DAZAP1 (DAZ associated protein1) (Naganuma et al 2012). Etonnamment, la protéine PSPC1 n'est pas indispensable à la formation des paraspeckles, et l'absence de son expression dans les cellules HeLa ne perturbe pas la formation normale des paraspeckles (Sasaki et al 2009).

## 2.2.3. Assemblage des paraspeckles

Par hybridation in situ de sondes oligonucléotidiques reconnaissant l'une ou l'autre des isoformes de Neat1, il a été possible de déterminer que Neat1-1 et Neat1-2 sont tous les deux présents au sein des paraspeckles (Naganuma et al 2012; Nakagawa et al 2011). Cependant les deux isoformes jouent des rôles différents dans les paraspeckles. La délétion de Neat1-2, et spécifiquement la perte de 10Kb de sa séquence en 3', induit une perte de la formation des paraspeckles (Sasaki et al 2009). En adéquation avec ces résultats, dans des fibroblastes embryonnaires provenant de souris KO Neat1, il a été observé que seule la surexpression de Neat1-2 restaure la formation des paraspeckles alors que la surexpression de Neat1-1 ne le permet pas (Naganuma et al 2012). A ce jour, Neat1-2 est donc considéré comme le facteur limitant dans la formation des paraspeckles et le rôle spécifique de Neat1-1 dans les paraspeckles n'a pas encore été élucidé.

L'assemblage des différentes protéines des paraspeckles avec Neat1 se fait au niveau du locus de Neat1 à proximité directe de la chromatine. Par des marquages du locus de Neat1, du lncRNA Neat1 et des protéines des paraspeckles PSPC1 et NONO, le décours temporel de la formation des paraspeckles entre la mitose et la formation d'une nouvelle cellule a été décrit. Dans le noyau de la cellule fille, les protéines PSPC1 et NONO sont présentes mais dispersées. Elles commencent à s'agglomérer après le début de la transcription de Neat1 (Clemson et al 2009). Grâce à l'imagerie cellulaire en temps réel, la cinétique de la fixation des protéines PSPC1, NONO et SFPQ sur le transcrit de Neat1 a pu être déterminée avec une précision de 5min permettant ainsi d'observer que les protéines du paraspeckles se fixent sur Neat1 au fur et à mesure de sa transcription (Mao et al 2011). La transcription de Neat1 est donc le déclencheur de la formation des paraspeckles. L'inhibition de la transcription de Neat1 par un bloqueur de l'ARN polymérase II empêche la formation et la maintenance des paraspeckles (Fox et al 2005). La présence de transcrits de Neat1 dans le noyau ne suffit pas à la formation de nouveaux paraspeckles (Mao et al 2011).

Le modèle communément accepté pour la formation des paraspeckles est donc que les protéines du paraspeckles se fixent sur Neat1-2 au fur et à mesure de sa formation. Si on ignore encore comment les protéines du paraspeckles sont spécifiquement recrutées, on sait qu'elle se lient à Neat1 par leurs motifs de reconnaissance des ARNs (RRM) permettant des liaisons ARN/protéines. Mais les protéines sont aussi dépendantes les unes des autres pour leur stabilisation au sein des paraspeckles. Par exemple, la localisation de PSPC1 dans les paraspeckles dépend de son domaine RRM et donc de son interaction avec Neat1, mais elle est aussi fortement dépendante de son interaction avec une autre protéine des paraspeckles, NONO, avec laquelle elle forme un hétérodimère (Fox et al 2005; Passon et al 2012). En plus de ce recrutement, il apparaît aussi que l'organisation entre elles des protéines du paraspeckle joue un rôle important dans la stabilisation des paraspeckles. Des études récentes montrent que SFPQ se polymérise en formant une super-hélice. Des mutations de cette super-hélice conduisent à une diminution de la présence de SFPQ dans les paraspeckles et une réduction du nombre de paraspeckles dans la cellule (Lee et al 2015). L'étude du rôle de la protéine HNRNPK sur la transcription de Neat1-2 est un autre exemple du rôle régulateur exercé par les protéines des paraspeckles sur leur formation. En permettant la transcription de l'isoforme Neat1-2, la protéine HNRNPK favorise la formation de nouveaux paraspeckles (Naganuma et al 2012). Cependant, la protéine HNRNPK étant elle même séquestrée dans les paraspeckles, elle ne peut activer la transcription de Neat1-2 que lorsque que le nombre de paraspeckles diminue et quelle se trouve à nouveau libérée dans le nucléoplasme. Ce mécanisme pourrait participer à la régulation du nombre de paraspeckles dans les cellules.

Une fois assemblés les paraspeckles présentent une organisation spatiale spécifique. En combinant immuno-marquage, hybridation in-situ et microscopie électronique dans des cellules HeLa et NIH3T3, il a été montré que la partie 5' de Neat1-1 est localisée en périphérie des paraspeckles au même titre que les extrémités 5' et 3' de Neat1-2, la séquence centrale de Neat1-2 se trouvant à l'intérieur des paraspeckles (Souquere et al 2010). Récemment un modèle d'organisation des paraspeckles a été proposé où les protéines seraient réparties entre une couche superficielle des paraspeckles (shell) et l'intérieur des paraspeckles (core) (Fig7.D) (West et al 2016). NONO, SFPQ, PSPC1 et FUS ont été trouvées dans la partie centrale des paraspeckles ; RBM14 est situé à l'interface entre la partie centrale et la périphérie ; d'autres protéines comme TDP-43 (transactive response DNA binding protein 43) sont retenues dans la couche superficielle des paraspeckles (West et al 2016).

## 2.3. Fonctions des paraspeckles

### 2.3.1. Rétention nucléaire des protéines

L'étude des domaines nucléaires organisés autour d'un lncRNA a montré que ces domaines pouvaient être capables de séquestrer des protéines nucléaires permettant ainsi l'inhibition de la fonction de ces protéines. Les granules nucléaires de stress en sont un exemple. Ces domaines nucléaires sont organisés autour du lncRNA Satellite III (SatIII) et sont impliqués dans la protection de la cellule vis-à-vis d'un stress de type choc thermique. En réaction à un choc, il a été observé une relocalisation de facteurs de transcription et de protéines d'épissage des ARNm vers les granules nucléaires de stress (Metz et al 2004). Il a été proposé que ce phénomène permettrait de modifier les gènes exprimés dans la cellule en favorisant l'expression de gènes de défense face à un stress cellulaire (Biamonti, & Vourc'h 2010).

Deux études publiées en 2014 suggèrent que la rétention de protéines dans les paraspeckles permettrait l'inhibition des fonctions de ces protéines dans le noyau. En induisant l'augmentation de la formation des paraspeckles par l'utilisation d'un inhibiteur du protéasome, ces équipes observent que la protéine SFPQ se trouve séquestrée en plus grande quantité dans les paraspeckles alors que sa localisation sur la chromatine diminue fortement (Imamura et al 2014). Cette relocalisation modifie l'expression des gènes cibles de SFPQ. Ainsi le gène Interleukine-8 (IL8) n'est plus réprimé par SFPQ et son expression augmente (Imamura et al 2014). A l'inverse, le gène ADARB2 (Adenosine deaminases acting on RNA B2) dont la transcription est régulée positivement par SFPQ, aura une expression réduite suite à l'augmentation de la rétention de SFPQ par les paraspeckles (Hirose et al 2013). Les 35 protéines présentes dans les paraspeckles identifiées en 2012 par Naganuma et al. dans les cellules HeLa (Naganuma et al 2012) sont majoritairement impliquées dans les mécanismes transcriptionnels ou post-transcriptionnels qui régulent l'expression de

gènes. Ainsi, les paraspeckles en retenant ces protéines peuvent inhiber leurs fonctions, et modifier par ce mécanisme l'expression de gènes régulés par ces protéines.

## 2.3.2. Rétentions d'ARNm

Dans les années 2000, plusieurs études permettent d'attribuer un autre rôle aux paraspeckles, celui de retenir des ARNs dans le novau des cellules où ils sont exprimés. En observant l'influence des mécanismes d'édition de l'Adénosine en Inosine (A-> I) sur la rétention nucléaire des ARN, une première étude montre par hybridation in situ et immuno-colocalisation que l'ARN mouse cationic amino acid transporter 2 (mCAT2) transcribed nuclear RNA (CTN-RNA) se situe dans les mêmes focis nucléaires que PSPC1 et NONO (Prasanth et al 2005) (Fig8). Cette colocalisation est observée dans la lignée NIH3T3 utilisée pour l'étude ainsi que dans des explants de foie de souris. Suite à cette découverte, plusieurs autres publications viennent confirmer la localisation d'autres ARNs dans les paraspeckles dans les lignées cellulaires HeLa, HEK293 et C2C12 (Chen et al 2008; Mao et al 2011). La démonstration que ces ARNs sont retenus dans le noyau par les paraspeckles a été réalisée grâce à l'inhibition de Neat1 par des oligonucléotides anti-Neat1 dans le modèle cellulaire HeLa. Suite à cette inhibition, les paraspeckles n'étant plus formés, une augmentation de l'export cytoplasmique d'ARNs qui étaient précédemment retenus dans le noyau est observée (Chen, & Carmichael 2009). La rétention d'ARNs est donc dépendante de Neat1 et de la formation des paraspeckles, mais elle est aussi dépendante de l'état des protéines présentes dans les paraspeckles. En effet, la reconnaissance des ARNs par les protéines du paraspeckle peut être inhibée, comme l'illustre l'étude de Hu et al., qui montre que la protéine CARM1 (coactivator-associated arginine methyltransferase 1) en méthylant la protéine NONO réduit son affinité pour les ARNs (Hu et al 2015; Elbarbary, & Maquat 2015).

Les paraspeckles reconnaissent et retiennent donc certains ARNs, en particulier ceux qui contiennent dans leur région 3' UTR un motif IRAlu qui est édité (édition des adénosines en inosines A->I). Les processus d'édition sont impliqués dans la régulation de l'expression des gènes à un niveau post-transcriptionnel. Lorsque cette édition a lieu sur une séquence codante d'un ARN, elle permet la diversification des protéines



# Figure 8: Rétention nucléaire par Neat1 et les paraspeckles d'ARNm hyper « édités » dans la région 3'UTR

Les ARNm présentant dans leur région 3' non codante (3'UTR) des séquences IRAlu (séquences répétées et inversées de type Alu) vont former des boucles d'ARN double-brin reconnues par l'enzyme ADAR (Adenosine deaminases acting on RNA) qui édite les bases Adénosines (A) en Inosines (I). Les ARNm contenant ces boucles d'ARN double brin, riches en inosines sont retenus à l'intérieur du noyau par les paraspeckles. L'expression de ces ARNm est alors inhibée par empêchement de leur translocation vers le cytoplasme.

qu'il code. Lorsque cette édition a lieu sur une région non codante d'un ARN, elle joue plutôt un rôle sur la stabilité de l'ARN en dirigeant cet ARN vers des voies de dégradation ou de rétention nucléaire. Dans l'ARN CTN ainsi que dans de nombreux autres ARNs, le phénomène d'édition se produit au niveau de motifs IRAlu de la région 3' UTR de l'ARN. La présence de motifs IRAlu ou de motifs IRSINE analogues des IRAlus chez les espèces non-primates permet la formation de boucles d'ARN double brin reconnues par la protéine ADAR (Adenosine deaminases acting on RNA) qui est responsable de l'édition A->I (Hundley, & Bass 2010; Liscovitch et al 2014). Avant la découverte des paraspeckles une étude avait déjà montré que des ARNs hyperédités en 3'UTR s'associaient avec les protéines NONO et SFPQ (Zhang, & Carmichael 2001) (Fig8). Suite à la caractérisation des paraspeckles, il a été montré que plusieurs ARNs endogènes exprimant ces éléments IRAlu dans leur 3'UTR colocalisent avec les différents éléments du paraspeckles (Chen et al 2008; Prasanth et al 2005; Mao et al 2011) . La séquence IRAlu éditée est donc un élément de reconnaissance par les paraspeckles. En effet, l'ajout de cette séquence dans la région 3'UTR d'un gène rapporteur comme l'egfp induit de manière significative la rétention nucléaire de l'ARNm de l'egfp ainsi que la diminution de l'expression de la protéine EGFP (Chen, & Carmichael 2008; Mao et al 2011) . Cependant, il a aussi été montré que des ARNs présentant des éléments IRAlu dans leur région 3'UTR sont localisés dans le cytoplasme de cellules humaines. Ces ARNs seraient concentrés dans des granules cytoplasmiques de stress impliqués dans l'inhibition de la traduction des ARNs en protéines (Fitzpatrick, & Huang 2012; Capshew et al 2012). Pour ces ARNs, la présence d'éléments IRAlus dans leur région 3'UTR n'est donc pas suffisante à leur rétention nucléaire par les paraspeckles.

D'autres motifs non identifiés à ce jour pourraient aussi être impliqués dans la reconnaissance des ARNs par les paraspeckles. Ces motifs peuvent aussi bien être des motifs de liaisons ARN/ARN reconnaissant le long ARN non codant Neat1, ou des motifs ARN/protéines reconnaissant les diverses protéines des paraspeckles. Du fait du grand nombre de protéines de liaison aux ARNs présentes dans les paraspeckles (Naganuma et al 2012), il est donc probable qu'il y ait autant de motifs de reconnaissance exprimés par les ARNs que de protéines dans les paraspeckles. De plus dans la séquence des ARNs, ces motifs pourraient être présents aussi bien dans la région 3'UTR, comme la majorité des études le montrent, que dans d'autres régions.

Une étude récemment publiée montre ainsi que l'hétérodimère formé par les protéines du paraspeckle NONO et SFPQ forme un site de reconnaissance permettant une liaison avec la région 5'UTR de certains ARNs (Duvignaud et al 2016).
# 3. Les fonctions physiologiques et pathologiques associées à Neat1 et aux paraspeckles

#### 3.1. Fonctions physiologiques de Neat1

De par les mécanismes moléculaires décrits pour les actions de Neat1 via les paraspeckles, on peut postuler une implication importante dans l'expression des gènes. Néanmoins les conséquences de ces régulations géniques sur les fonctions physiologiques d'un organisme sont très peu documentées. Un nombre limité d'études permet cependant d'attribuer certains rôles physiologiques à Neat1 et aux paraspeckles.

Le premier modèle de souris dans lequel l'expression de Neat1 est génétiquement inhibée est rapporté en 2011. Dans ces souris, un signal de polyadénylation a été ajouté juste après le début du site de transcription de Neat1. Cette mutation réduit presque totalement l'expression de Neat1-1 et abolit complétement l'expression de Neat1-2. La perte de formation des paraspeckles est confirmée après immunomarquage de PSPC1 qui n'est plus retrouvée agrégée dans des focis nucléaires mais est exprimée de manière dispersée dans le noyau (Nakagawa et al 2011). Dans une première analyse de ce modèle, les auteurs rapportent que ces souris mutées sont capables de se reproduire et que l'absence bi-allélique de Neat1 n'est pas létale suggérant que Neat1 n'est pas indispensable dans l'organisme. Les animaux sont viables, fertiles et leur masse corporelle est normale à l'âge adulte. En particulier l'épithélium gastrique qui est le tissu dans lequel Neat1 s'exprime le plus fortement, n'a pas une organisation histologique modifiée par la perte de l'expression de Neat1 (Nakagawa et al 2011).

Pourtant, en 2014 deux études réalisées sur ce même modèle de souris KO-Neat1 montrent que la perte de Neat1 a des conséquences sur certaines fonctions physiologiques.

L'une de ces études vient préciser l'influence de Neat1 sur la fertilité des souris. Les souris KO-Neat1 malgré une ovulation normale ont un taux de grossesse réduit par rapport à des souris WT (Nakagawa et al 2014). En effet la moitié des souris Neat1 KO échouent dans le développement d'une grossesse en raison d'un mauvais développement du corps jaune dû à une absence d'élévation des taux sériques de progestérone au moment de la mise en place de la grossesse. Cette augmentation de progestérone sécrétée par le corps jaune est indispensable à la nidation de l'embryon dans le placenta, et son absence cause une augmentation du nombre d'échecs de grossesse. L'injection de progestérone chez ces souris restaure partiellement la fertilité des femelles. Neat1 semble exercer son rôle directement dans les cellules lutéales du corps jaune. En effet, la greffe d'ovaires de souris WT restaure aussi partiellement la fécondité des souris, suggérant une action locale de Neat1 sur la fonction lutéale. Cette hypothèse est confortée par l'observation de modifications de l'organisation histologique des cellules lutéales de souris Neat1 KO, qui met en évidence un cytoplasme de taille réduite et des espaces vésiculaires inter-cellulaires. Au niveau de l'expression des gènes, la perte d'expression de Neat1 modifie l'expression de gènes clefs pour la fonction lutéale. Sont ainsi observées une diminution des gènes requis pour la génération (Lhcgr) et le fonctionnement (Star, Hsd3b) du corps jaune et une augmentation d'un gène qui inhibe la formation du corps jaune (Akr1c18) (Nakagawa et al 2014).

La seconde étude publiée sur le modèle de souris Neat1 KO montre un rôle de Neat1 dans la formation de la glande mammaire (Standaert et al 2014). Cette étude montre au départ que dans des portées issues de mères homozygotes Neat1 KO, seulement 24% des nouveaux nés survivent. En comparaison avec des petits issus de mères hétérozygotes ou WT, ces nouveaux nés sont plus petits et plus maigres. La raison de ce défaut de croissance provient d'une insuffisance d'allaitement par la mère et peut être restauré si les petits sont donnés en nourrice à une mère WT. L'observation histologique de la glande mammaire chez les souris montre que chez les femelles gestantes et allaitantes WT, Neat1 est exprimé dans l'épithélium ductal de la glande mammaire, où il forme des paraspeckles. En revanche chez les souris homozygotes KO pour Neat1, il y a une modification de la structure de l'épithélium mammaire. Le nombre d'alvéoles mais aussi le nombre de branchements et la longueur des canaux sont réduits. La surface de l'épithélium est considérablement réduite par une altération non pas de la taille des cellules mammaires mais de leur prolifération durant la gestation. En revanche, les protéines sécrétées dans le lait ne semblent pas impactées par l'absence de Neat1. Il semblerait donc que Neat1 joue un rôle important dans le développement de la glande mammaire en condition de lactation mais pas dans sa fonction sécrétrice. Cependant cette étude ne permet pas de déterminer si cette altération du développement de la glande mammaire est la conséquence de la perte locale de l'expression de Neat1 ou un effet lié à la perte de Neat1 dans le reste de l'organisme.

En plus des fonctions de Neat1 identifiées grâce à l'étude du modèle de souris KO, d'autres travaux semblent suggérer son implication dans la différenciation et la maturation des cellules. Il a été en effet observé que les rares tissus n'exprimant pas Neat1 sont des tissus embryonnaires (Nakagawa et al 2011). Différentes équipes ont rapporté l'augmentation de l'expression de Neat1 et donc des paraspeckles au cours de la différenciation et de la maturation dans plusieurs modèles cellulaires. Ainsi, les cellules embryonnaires humaines (hESCs) expriment les principales protéines des paraspeckles (NONO, SFPQ, PSPC1) mais ces protéines ne s'agglomèrent pas en paraspeckles et ce n'est qu'après la différenciation de ces cellules en trophoblastes que Neat1 est exprimé et que les paraspeckles se forment (Chen, & Carmichael 2009). Une observation similaire est faite dans le processus de développement musculaire au cours duquel les myoblastes se différencient en myotubes. Durant ce processus, la quantité de Neat1 triple (Sunwoo et al 2009), tandis que le taux de l'isoforme Neat1-2 peut être multiplié par six (Lehnert et al 2007). Cette augmentation de l'expression de Neat1 semble induire le grossissement des paraspeckles (Sunwoo et al 2009). Enfin, dans des oligodendrocytes matures ou dans des cellules GABAergiques corticales, l'expression de Neat1 double en moyenne par rapport aux cellules souches neurales dont ces deux types cellulaires sont issus (Mercer et al 2010).

Ces études ont permis d'établir un lien corrélatif entre la différenciation et la maturation cellulaire et le niveau d'expression de Neat1 et la formation des paraspeckles. Cependant cette corrélation n'a été testée fonctionnellement que dans une seule étude portant sur la différenciation des cellules de moelle osseuse. Dans ces cellules, la différenciation induite par le trans-acide rétinoïque (ATRA) est associée à

une augmentation de l'expression de Neat1. L'utilisation de siRNA dirigés contre Neat1 réduit la différenciation induite par ATRA dans ces cellules (Zeng et al 2014). Ces résultats suggèrent que Neat1 et les paraspeckles pourraient jouer un rôle actif dans les processus de différenciation.

Les données concernant les rôles physiologiques de Neat1 et des paraspeckles sont encore parcellaires. Il apparaît néanmoins que l'abolition de Neat1 et des paraspeckles n'induit pas des perturbations de l'organisme suffisamment graves pour être létales, bien que plusieurs fonctions vitales de reproduction et de développement soient affectées au moins partiellement. Le développement et l'étude de différents modèles animaux et cellulaires devraient permettre dans le futur d'élucider l'ensemble des fonctions physiologiques impliquant Neat1 et les paraspeckles.

#### 3.2. Les pathologies associées à Neat1

De très nombreuses études rapportent des modifications de l'expression de Neat1 et de gènes contrôlés par Neat1 dans des contextes pathologiques. En particulier, dans une étude comparant l'expression de lncRNAs dans des tissus sains et des tissus cancéreux humains, il a été observé que Neat1 était, parmi les lncRNAs, celui dont l'expression est la plus altérée en cas de cancer (Gibb et al 2011). Sur 17 types de cancers analysés, Neat1 varie de manière importante dans 14 cas (cerveau, poumons, moelle osseuse, muscles, œsophage, estomac, colon, vessie, ganglions lymphatiques, thyroïde, seins, ovaires, prostate, testicules). Plusieurs études montrent dans des cancers cérébraux, du sein, des poumons, de l'œsophage, de l'estomac, du foie, des ovaires et de la prostate qu'une augmentation de Neat1 est corrélée à un mauvais pronostic vital, à une augmentation de la taille des tumeurs et de leurs métastases, et à une augmentation de leur stade de progression (données compilées dans la revue (Lo et al 2016)).

Il existe de nombreuses autres pathologies associées à une surexpression de Neat1. Ainsi différentes infections virales entraînent une surexpression de Neat1 comme cela a été montré chez la souris après infection par le virus encéphalite (Saha et al 2006) ou par le virus du VIH (Zhang et al 2013). L'étude de Zhang et al. suggère que dans le cas de l'infection par le virus du VIH, Neat1 aurait plutôt un rôle protecteur qui limiterait la reproduction du virus dans l'organisme (Zhang et al 2013).

Dans la sclérose latérale amyotrophique (SLA), plusieurs études rapportent l'augmentation de l'expression de Neat1 dans les stades précoces de la maladie (Nishimoto et al 2013; Ito et al 2011). Neat1 et les paraspeckles pourraient jouer un rôle important dans cette pathologie car plusieurs protéines (dont FUS et TDP-43) fortement impliquées dans le développement de la SLA sont des protéines des paraspeckles (Lourenco et al 2015).

En revanche dans certaines pathologies, une réduction de l'expression de Neat1 a été rapportéecomme dans la leucémie myéloïde aigüe (Zeng et al 2014). Dans cette pathologie cancéreuse contrairement à la plupart des autres cancers, Neat1 n'aurait pas un rôle oncogène mais plutôt suppresseur de tumeur. Chez les patients souffrant de la dystrophie musculaire de Duchenne, l'expression de Neat1 est également diminuée (Timmons et al 2005).

Il apparaît donc que de nombreuses situations en pathologie humaine sont associées à des défauts d'expression de Neat1. Il reste cependant à déterminer si ces défauts d'expression peuvent être seulement considérés comme des marqueurs diagnostiques ou si elles participent activement au développement de la pathologie et constituent de ce fait une cible thérapeutique potentielle.

Résultats expérimentaux

# A RNA pull down procedure to indentify RNA targets of a long noncoding RNA

Manon Torres, Denis Becquet, Séverine Guillen, Bénédicte Boyer, Mathias Moreno, Marie-Pierre Blanchard1, Jean-Louis Franc and Anne-Marie François-Bellan

#### Introduction

Les longs ARN non codants (lncRNA) se définissent par leur longueur supérieure à 200 nucléotides et leur absence de cadre de lecture. Le nombre de ces lncRNAs chez les mammifères n'a cessé d'augmenter et est aujourd'hui estimé chez l'homme à plus de 10,000 transcrits, et pour certains d'entre eux leur implication dans d'importants circuits de régulation de l'expression des gènes a été identifiée (Redrup et al 2009; ENCODE Project Consortium 2012). Même si leurs fonctions biologiques demeurent largement méconnues, on sait à partir d'études spécifiques de certains lncRNAs qu'ils peuvent être impliqués dans la régulation de l'expression des gènes par des mécanismes très divers pouvant impliquer soit un contrôle épigénétique, soit un contrôle de la transcription ou encore un contrôle post-transcriptionnel de la maturation des ARN. Dans ce dernier cas, le contrôle exercé par le lncRNA peut être direct ou indirect à travers une association avec des protéines de liaison aux ARNs. Malat1 et Neat1 sont des lncRNAs de localisation nucléaire qui sont parmi les lncRNAs les plus exprimés chez les mammifères. Ils sont impliqués dans le contrôle de la maturation des ARNs au travers de deux mécanismes distincts. Malat1 localisé dans des sous-domaines nucléaires appelés speckles s'associe à des protéines d'épissage des ARNm et participe à la régulation de l'épissage de nombreux ARNs immatures (Engreitz et al 2014). Neat1 est localisé dans des sous-domaines nucléaires voisins des speckles, les paraspeckles, dans lesquels il est associé à plusieurs protéines de liaison aux ARNs. Les paraspeckles peuvent séquestrer dans le noyau certains ARNs et ainsi inhiber leur export cytoplasmique. Ces deux lncRNAs sont donc impliqués au travers de mécanismes différents dans le contrôle de l'expression d'ARNs dont on ne connaît ni le nombre ni l'identité.

C'est pour cette raison que nous avons développé une méthode de RNA pull-down avec pour objectif de pouvoir identifier l'intégralité des cibles ARNs d'un lncRNA. La finalité de cette méthode est d'isoler spécifiquement un lncRNA à partir d'un échantillon cellulaire ou tissulaire, et d'entrainer avec ce lncRNA l'ensemble des ARNs qui lui sont associés. Les conditions expérimentales de la méthode ont été optimisées afin de proposer un protocole efficace et adaptable à différents lncRNAs ainsi qu'à différents modèles expérimentaux, cellulaires ou tissulaires..

#### **Résultats**

Le RNA pull-down s'inspire de méthodes publiées sous le nom de ChIRP (Chu et al 2011; Chu et al 2012) ou CHART (Simon et al 2011; Simon 2013), qui permettent d'isoler la chromatine associée à un lncRNA. Le protocole a été repensé pour permettre l'extraction des ARNs associés à un lncRNA.

Deux étapes sont cruciales pour la bonne réalisation d'un RNA pull-down. D'une part la conception de sondes oligonucléotidiques spécifiques du lncRNA d'intérêt et d'autre part la préparation adéquate du matériel cellulaire ou tissulaire utilisé. Ces deux étapes doivent être ajustées en fonction particulièrement des caractéristiques du lncRNA ciblé et du modèle expérimental cellulaire ou tissulaire étudié.

La conception des sondes oligonucléotidiques est une étape cruciale qui doit assurer à ces sondes une excellente affinité et spécificité vis-à-vis du lncRNA d'intérêt. Nous proposons une stratégie de conception des sondes qui se base sur une analyse bioinformatique des régions du lncRNA accessibles à l'hybridation par les sondes. Nous avons ainsi conçu plusieurs sondes que nous avons ultérieurement testées expérimentalement. Ces sondes s'étant révélées aussi efficaces individuellement qu'en cocktail, elles peuvent être utilisées séparément.

La préparation du matériel biologique, cellulaire ou tissulaire est l'étape nécessitant le plus de mise au point. Nous avons utilisé un protocole de préparation du matériel qui permet de réaliser l'étude aussi bien « in vitro » dans des lignées cellulaires en culture (entre 10<sup>6</sup> et 10<sup>7</sup> cellules) qu' « ex vivo » à partir de prélèvements de tissus (entre 1 et 10mg). La préparation nécessite en premier lieu une fixation (ou Cross-linking) du matériel afin d'assurer la stabilité des interactions entre le lncRNA d'intérêt et ses cibles ARNs, et d'autre part un traitement par ultrasons des échantillons permettant de fragmenter l'ADN qui entraîne une viscosité de l'échantillon empêchant ainsi sa manipulation. Nous proposons aussi différentes stratégies qui permettent d'ajuster les conditions expérimentales et de réaliser au mieux ces étapes dans différents modèles d'étude.

A la fin du protocole, les échantillons contenant l'intégralité des ARNs associés au lncRNA d'intérêt peuvent être étudié par RT-qPCR dans le cadre d'une recherche

ciblée, ou par séquençage haut débit des ARNs afin de déterminer la liste exhaustive des ARNs associés à un lncRNA particulier.

Nous avons validé notre méthode de RNA pull-down par des exemples montrant l'efficacité de sondes spécifiques conçues selon notre stratégie pour obtenir un fort enrichissement en Neat1 mais aussi en un autre lncRNA, Malat1. Nous montrons aussi que les ARNs de l'hormone de croissance (Gh) et de la prolactine (Prl) sont associés au lncRNA Neat1 aussi bien dans le modèle cellulaire GH4C1 que dans des prélèvements de tissus hypophysaires de souris. Enfin, nous avons pu montrer un enrichissement réciproque de Malat1 par Neat1 RNA pull-down et réciproquement de Neat1 par Malat1 pull-down dans les cellules GH4C1. Finalement, le RNA pull-down de Neat1 couplé à du séquençage haut débit des ARNs nous a permis d'obtenir la liste complète des ARNs associés à Neat1 dans les cellules GH4C1.

### Conclusion

Le développement d'outils permettant l'étude des interactions entre des lncRNAs et des ARNs est un enjeu majeur pour la compréhension du rôle de ces lncRNAs dans le fonctionnement cellulaire. Afin de proposer une méthode performante et adaptable à différentes conditions expérimentales, le RNA pull-down a été développé et optimisé pour être utilisable pour l'étude de n'importe quel lncRNA et cela aussi bien dans des explants de tissus que dans des cellules en culture. Il permet ainsi en quelques jours d'isoler l'intégralité des cibles ARNs d'un lncRNA qui peuvent ensuite être analysées par RT-qPCR ou, plus exhaustivement, par séquençage haut débit des ARNs. Soumis au Journal of Visualized Experiments (JoVE), le protocole de RNA pull-down décrit en détails les différentes étapes de la réalisation de la méthode. Accompagné d'un tutoriel vidéo, ce support permet à tout expérimentateur d'être guidé dans la reproduction du RNA pull-down.

# TITLE: A RNA pull-down procedure to identify RNA targets of a long noncoding RNA

# **AUTHORS & AFFILIATIONS:**

Manon Torres, Denis Becquet, Séverine Guillen, Bénédicte Boyer, Mathias Moreno, Marie-Pierre Blanchard<sup>1</sup>, Jean-Louis Franc and Anne-Marie François-Bellan

Aix-Marseille Université, CNRS, CRN2M-UMR7286, Faculté de Médecine Nord, 51 Bd

Pierre Dramard CS 80011, 13344 Marseille cedex 15, France

<sup>1</sup> Present address: Plate-forme d'imagerie MRI, UMS3426 Institut de Génétique Humaine-CNRS 141, rue de la Cardonille, 34396 Montpellier cedex 5, France

**KEYWORDS:** long noncoding RNA (lnc RNA), RNA pull-down, probe design, predicted RNA secondary structure, network cross-linking, DNA and RNA shearing

#### **SHORT ABSTRACT:** (10 word minimum, 50 word maximum)

This RNA pull-down method allows identifying the RNA targets of a long non-coding RNA (lncRNA). Based on the hybridization of home-made designed anti-sense DNA oligonucleotide probes specific to this lncRNA in an appropriately fixed tissue or cell line, it efficiently allows the capture of all RNA targets of the lncRNA.

#### LONG ABSTRACT: (150 word minimum, 300 word maximum)

Long non-coding RNAs (IncRNA) that correspond to sequences of more than 200 nucleotides without a defined reading frame, belong to the regulatory non-coding RNA's family. Although their biological functions remain largely unknown, the number of these IncRNAs has steadily increased and is now estimated in humans at more than 10,000 transcripts. Some of them are known to be involved in important regulatory pathways of gene expression taking place at the transcriptional level but also at different steps of RNA co-and post-transcriptional maturation. In the latter cases, RNAs that are targeted by the IncRNA have to be identified. That's the reason why it is useful to develop a method enabling the identification of RNAs associated directly or indirectly with a IncRNA of interest.

This protocol inspired by previously published protocols allowing the isolation of a IncRNA together with its associated chromatin sequences, was adapted to permit the isolation of associated RNAs. We determined that two steps are critical for the efficiency of this

protocol. The first one is the design of specific anti-sense DNA oligonucleotide probes able to hybridize to the lncRNA of interest. To this end, the lncRNA secondary structure was predicted by bioinformatics and anti-sense oligonucleotide probes with a strong affinity for regions that display a low probability of internal base pairing were designed. The second crucial step of the procedure relies on the fixative conditions of the tissue or cultured cells that have to preserve the network between all molecular partners. Coupled with high throughput RNA sequencing, this RNA pull-down protocol can provide the whole RNA interactome of a lncRNA of interest.

#### **INTRODUCTION:**

The overall goal of the method described here is to identify RNA molecular partners of a long noncoding RNA (lncRNA). LncRNA correspond to sequences of more than 200 nucleotides without a defined reading frame. Some of them have been shown to be involved in gene expression regulation not only at the transcriptional level but also at different steps of RNA co-and post-transcriptional maturation. In the latter cases, molecular partners of the lncRNA are RNAs that are to be identified. A method enabling the identification of RNAs associated directly or indirectly with a lncRNA of interest is then essential to be developed.

Previously published methods such as Chromatin Isolation by RNA Purification (ChIRP) allow high-throughput discovery of RNA-bound proteins and DNA binding sites (Chu et al 2011; Chu et al 2012). The IncRNA of interest is hybridized to biotinylated complementary oligonucleotides, and isolated using streptavidin beads. On the other hand, Simon et al. have developed a method termed Capture Hybridization Analysis of RNA Targets (CHART) to determine the genomic binding sites of a specific IncRNA (Simon et al 2011; Simon 2013). CHART is also a hybridization-based technique that is used to specifically enrich the protein and DNA targets of endogenous IncRNAs. This method is similar to ChIRP; however, the main difference lies in the design of probes to target IncRNAs. ChIRP has been inspired by RNA FISH to design a pool of short complementary DNA oligonucleotide probes to tile the entire length of the IncRNA. On the other hand, Simon et al. have adapted an RNaseH mapping assay to probe sites on IncRNAs that are available to hybridization. The procedure proposed here to design the anti-sense DNA biotinylated oligonucleotide probes uses bioinformatics modeling of IncRNA secondary structure (Reuter, & Mathews 2010), to select probes with a strong affinity for regions that display a low probability of internal base

pairing. This procedure has the advantage to be less expensive than that based on pools of tiling oligonucleotide probes (Chu et al 2012) and less time consuming than that based on RNAse-H sensitivity (Simon 2013).

As there is a growing body of evidence for a post-transcriptional gene regulation by lncRNAs (Sun et al 2017), it is very useful to develop an approach enabling to capture RNAs that are targets of a lncRNA. Furthermore, to be usable for most applications that people can need, the approach was optimized both in cultured cells and in tissue extracts.

# **PROTOCOL:**

#### 1. Probe design

1.1. Using the primary sequence of the lncRNA of interest, its secondary structure can be generated on the "RNAstructure Webserver" (Reuter, & Mathews 2010). On this site different algorithms can be used. The three prediction tools that give the best results for the probe designs are: Fold (lowest free energy structure), MaxExpect (highly probable base pairs) and Probnot (probable base pairs including pseudoknots). These three analyses can be performed and compared. Other web servers such as the Vienna RNA web suite (Gruber et al 2008) can also be used.

Select regions that display a low probability of internal base pairing and design antisense oligonucleotide probes of 25 bases with a strong affinity for these regions. The GC contents of these probes should be comprised between 40 and 60 %. Make sure, using an alignment search tool (Blast), that the selected anti-sense oligonucleotide probes do not recognize nucleotide sequences in other RNA expressed in your cell system.

- 1.2. Design also a non-specific DNA oligonucleotide probes of 25 bases which displays neither affinity for the lncRNA of interest nor for other RNA sequences in your genome of interest.
- 1.3. Order the probes with a biotin at 3'-end. The distance between oligonucleotide and Biotin has to be increased with a triethyleneglycerol spacer.

NOTE: For an optimal result and to be able to assess the specificity of results of the pull-

down, it is recommended to design 3 different anti-sense oligonucleotide probes and then to compare experimentally their efficiency.

# 2. Cross-linking

# 2.1. Cultured cell cross-linking

- 2.1.1. Culture the GH4C1 sommatolactotroph pituitary cells in HamF10 medium supplemented with 15% horse serum and 2% fetal calf serum. The cells must be grown until confluence in 78.5 cm<sup>2</sup> culture plates. This corresponds approximately to 1.10<sup>7</sup> cells.
- 2.1.2. Remove the cell culture medium from a confluent GH4C1 78.5 cm<sup>2</sup> culture plate then rinse with 1X the medium volume of phosphate buffered saline (PBS)
- 2.1.3. Fix the cells with a 1% paraformaldehyde solution in PBS (10ml for a 78.5 cm<sup>2</sup> dishes); this solution must be freshly prepared from a 4% paraformaldehyde stock solution. Cross-link under agitation for 10 min at room temperature (RT).
- 2.1.4. Quench the paraformaldehyde action by adding 1/10 volume of glycine1,25M (1 ml per 10ml of paraformaldehyde solution); agitate 5 min at RT.
- 2.1.5. Discard the media by aspiration and rinse two times (5 min) with 1X the medium volume of PBS
- 2.1.6. Add a volume of PBS corresponding to 1/10 of your volume of media and collect your cells with a cell scraper and transfer to an Eppendorf tube.
- 2.1.7. Spin at 510 rcf at 4°C for 5 min.
- 2.1.8. Remove as much supernatant as possible.
- 2.1.9. Store pellets indefinitely at -80°C.

# 2.2. Tissue cross-linking

2.2.1. Put your fresh tissue in a solution of 1% paraformaldehyde diluted in PBS (approximately 10X the volume of the tissue), agitate for 10 min at RT.

- 2.2.2. Quench the paraformaldehyde action by adding a 1.25M glycine solution (1 ml per 10ml of paraformaldehyde solution) and agitate 5 min at RT.
- 2.2.3. Discard by aspiration the media and rinse two times with PBS (approximately 10X the volume of the tissue).
- 2.2.4. Remove as much supernatant as possible.
- 2.2.5. Store the cross-linked tissue indefinitely at -80°C.

# 3. Cell or tissue lysis

- 3.1. Prepare the Lysis Buffer (50mM Tris-HCl pH 7.0, 10mM EDTA, 1% SDS) supplemented with 200 U/ml of a RNAse inhibitor solution and a cocktail of proteases inhibitor 5 μl/ml (Sigma)).
- 3.2. To obtain lysed samples, re-suspend the cell pellets or cross-linked tissues with this buffer (approximately 1ml per 100mg of cell pellet or tissue). A cell pellet obtained from 1.10<sup>7</sup> cells give rise roughly to a lysed sample containing 20 mg of proteins.

NOTE: Depending on the tissue used, a step of mechanical disruption should be added. In this case, it is important to avoid heating of the samples during this additional step.

# 4. Sonication

#### 4.1. Optimization of the sonication conditions:

- 4.1.1. Program the Bioruptor with 5 to 10 series of 30 sec ON and 30 sec OFF.
- 4.1.2. Tests to optimize sonication conditions can be performed on diluted lysed samples (dilution factor ½ or ¼ corresponding to approximately 10 or 5 mg of protein). Place diluted lysed samples in the 4°C water bath and start the sonication series.
- 4.1.3. Extract RNAs either with a RNA purification kit or with Trizol.
- 4.1.4. Load the totality of purified RNA on a 1% of agarose gel electrophoresis in TBE buffer, check the length of RNA fragments. This length should range between 200 and 800 bp.

NOTE: Depending on the RNA fragment size, the anti-sense oligonucleotide probes efficiency can vary. It is then recommended to check the efficiency of the probes under different conditions of sonication.

#### 4.2. Sonication of the lysed samples:

- 4.2.1. Place lysed samples containing 20 mg of proteins (obtained after step 3.2) in the 4°C water bath and start the sonication series as optimized above.
- 4.2.2. Immediately after sonication, centrifuge 5 min at 12,000 g at 4°C. Transfer supernatants in new Eppendorf tubes. To ensure homogeneity, replicate supernatants can be pooled and redistributed at this step.

#### 5. RNA pull-down

#### 5.1. Day 1 – Hybridization step

- 5.1.1. Add 2 volumes of hybridization buffer (50mM Tris-HCl pH 7.0, 750 mM NaCl, 1 mM EDTA, 1% SDS, 15% Formamide added extemporaneously) to supernatants collected after the sonication step. Vortex.
- 5.1.2. Transfer 20  $\mu$ l of each sample in an Eppendorf tube (input samples) and store at -20°C.
- 5.1.3. Add 100 pmol of biotinylated oligonucleotide probes (specific or non-specific; see Table 1)) to each sample. Incubate 4 to 6 h under moderate agitation on an Eppendorf rotator at RT.
- 5.1.4. Add 50  $\mu$ l of magnetic streptavidin beads supplemented with 200 U/ml of a RNAse inhibitor solution and a cocktail of proteases inhibitor 5  $\mu$ l/ml.
- 5.1.5. Incubate overnight at 4°C under moderate agitation on an Eppendorf rotator at RT.

#### 5.2. Day 2 – RNA isolation step

- 5.2.1. Use magnetic support to separate beads from cell lysate, discard supernatant and wash the beads with 900  $\mu$ l of wash buffer (SDS 0.5%, SSC 2X). Repeat 5 times interspersed with 5 min agitation on the rotator at RT.
- 5.2.2. After the last wash, decant a last time and add in the samples 95 μl of Proteinease K buffer (10 mM Tris-HCl pH 7.0, 100 mM NaCl, 1 mM EDTA, 0.5% SDS) and 5μl of proteinase-K (20mg/ml).
- 5.2.3. Defrost the input samples and add 75  $\mu$ l of Proteinease K buffer and 5  $\mu$ l of proteinase-K (20mg/ml).
- 5.2.4. Incubate samples and input samples 45min at 50°C then 10 min at 95°C.
- 5.2.5. Chill sample on ice before separating the beads from RNAs with the magnetic support. Keep the supernatant and discard the beads.
- 5.2.6. Purify RNAs with a RNA purification kit and include a DNA digestion step. Store RNAs at -80°C.

#### **REPRESENTATIVE RESULTS:**

Several recent studies have shown that lncRNAs play essential role in almost every important biological process and that this role is achieved through the control of gene expression occurring both at the transcriptional and the post-transcriptional levels showing in this latter case that RNAs may be the target of lncRNAs (Sun et al 2017).

The lncRNA Nuclear enriched abundant transcript 1 (Neat1) is implicated in different neuropathologies as frontotemporal dementia, amyotrophic lateral sclerosis or epilepsy (Tollervey et al 2011; Riva et al 2016; Barry et al 2017) and is also misregulated in different cancers (Adriaens et al 2016; Fang et al 2017).

This lncRNA is also known to be the structural component of specific nuclear bodies, the paraspeckles, and to be involved in post-transcriptional circadian regulation of gene expression (Torres et al 2016). Paraspeckles that are found in every cell nucleus and are formed around not only Neat1 that is necessary for their formation, but also around several RNA binding proteins (RBP) (Fox, & Lamond 2010), are indeed known to be able to retain RNA targets within the nucleus (Chen et al 2008). The formation of paraspeckles is achieved through the association of the different components. This formation was shown to display a circadian rhythmic pattern driving a rhythmic nuclear retention of RNA targets (Torres et al 2016). The nuclear retention of RNA targets by paraspeckles may occur through binding to RBP or directly through RNA/RNA association, but the extend of RNAs targeted by paraspeckles had to be determined. To identify the RNA targeted directly or indirectly by Neat1, a RNA pull-down protocol was designed that allows the isolation and the identification of all Neat1 RNA targets in cultured cells as well as in tissue samples (see Fig.1 for a graphical presentation of the technique).

The protocol was also successfully applied to the identification of RNA targets of an another lncRNA Metastasis Associated Lung Adenocarcinoma Transcript 1 (Malat1). Malat1 is a highly conserved and expressed lncRNA found in nuclear speckles together with several RNA splicing factors. Malat1 is known to be involved in the regulation of the splicing of several nascent pre-mRNA (Tripathi et al 2010; Engreitz et al 2014).

Specific (SO) and non-specific oligonucleotide (NSO) probes were generated using the probe design strategy described here. This strategy relies on the selection of regions that display a low probability of internal base pairing as predicted by the secondary structure of the lncRNA and on the design of specific probes with a strong affinity for these regions. As a representative result of these bioinformatics predictions, a picture of the predicted secondary structure of a sequence of Neat1 (nucleotides 1480 to 2000) together with the position of two designed SO probes are given in Fig.2.

The designed probes were directed to rat Neat1 or Malat1 for GH4C1 cultured cells and to mouse Neat1 for pituitary tissue extracts (Table 1). The relative enrichment in Neat1 or Malat1 was calculated for non-specific and specific probes relative to the input samples. Figure 3 shows the efficiency of the specific probes to pull-down Neat1 in rat GH4C1 pituitary cell line (Fig3A) and in mouse pituitary tissue extracts (Fig3B). When using the probe design protocol to generate specific oligonucleotide (SO) probes

directed to Malat1, one efficient probe was obtained while another was not efficient enough and was discarded (Fig5A).

After a RNA pull-down procedure followed by RT-qPCR experiments, some RNAs assessed with specific primers (Table1) were shown to be associated with Neat1 or Malat1 in GH4C1 extracts. The RNAs associated with Neat1 in GH4C1 cell extracts were also shown to be associated with Neat1 in pituitary tissue extracts. Indeed, after Neat1 RNA pull-down, Malat1 was found to be targeted by Neat1 both in GH4C1 cell line and in pituitary mouse tissue extracts (Fig4A). Reciprocally, Neat1 was significantly enriched after Malat1 RNA pull-down performed with a specific probe in GH4C1 cells (Fig5B). These results by highlighting the close relationship between the two lncRNAs are consistent with the potential co-regulatory role of Neat1 and Malat1 suggested by Malat1 knockout mice that display variations in Neat1 RNA expression (Nakagawa et al 2012; Zhang et al 2012). The transcripts of two main pituitary hormones, growth hormone (Gh) (Fig4B) and prolactin (Prl) (Fig4C) were significantly enriched following Neat1 RNA pull-down with specific probes in both GH4C1 cells and pituitary extracts suggesting a possible regulation of the two hormones by Neat1. When comparing the two specific probes used, it appeared that their efficiency could vary depending on the RNA target considered (Fig4B and Fig4C). These results highlight the necessity to design several specific probes in order to select those displaying not only the best efficiency in the enrichment of the pull-down lncRNA but also the best efficiency in the enrichment of its RNA targets.

The RNA pull-down method can also be followed by RNA high-throughput sequencing to obtain the comprehensive list of RNA targets of a lncRNA of interest. A RNA-seq analysis on GH4C1 pituitary cells after Neat1 RNA pull-down using the two specific probes described above was performed. It should be noted that a negative control using a NSO could also be subjected to RNA-seq analysis if level of RNA recovered after the RNA pull-down with NSO is sufficient to allow the construction of libraries. This was not the case in our hands. Libraries that were generated after use of specific probes were analyzed using Tophat/cufflinks pipeline (Trapnell et al 2012) and only transcripts with values of fragment per kilobase per million of mapped reads (FPKM) higher than 1 were taken into account. The lists obtained with the two specific probes were crossed to assess the specificity of the results. 4268 genes were found associated with paraspeckles, which represented 28% of expressed transcripts in GH4C1 cells (Torres et al 2016). Consistent with results obtained using qPCR analysis the transcripts of Gh, Prl and Malat1 were found to be associated with Neat1. The RNA pull-down method has therefore been proved to be an efficient tool to explore the interaction between lncRNAs and their RNA-targets.

# FIGURE AND TABLE LEGENDS:

Figure1: Graphical representation of the RNA pull down procedure

Figure 2: Secondary structure of a Neat1 sequence (nucleotide 1480 to 2000) as predicted by bioinformatics resource (the RNAstructure webserver; lowest free energy structure). The structure is colored according to degree probability of base pairing. The two oligonucleotides probes (SO1 and SO2) in red are positioned along the Neat1 RNA structure.

Figure3: qPCR validation of Neat1 enrichment versus input after Neat1 RNA pull down by two different specific probes (SO1-rn and SO2-rn for GH4C1 cells and SO1-mm and SO2-mm for pituitary tissue) as compared to a non-specific one (NSO-rn for GH4C1 cells and NSO-mm for pituitary tissue) in GH4C1 rat cells (A) and in mouse pituitary tissue extracts (B). Results are mean± SEM obtained in 3 to 10 experiments.

Figure4: qPCR validation of Malat1 (A), Gh (B), Prl (C) enrichment versus input after Neat1 RNA pull down using different specific probes as compared to a non-specific one in GH4C1 rat cells and mouse pituitary tissue extracts. Results are mean± SEM obtained in 3 to 8 experiments.

Figure5: qPCR validation of Malat1 (A) and Neat1 (B) enrichment versus input after Malat1 RNA pull down by two different specific probes (SO3-rn and SO4-rn) as compared to a non-specific one (NSO-rn) in GH4C1 rat cells. Results are mean± SEM obtained in 3 experiments.

Table 1: Sequences of DNA oligonucleotide probes and qPCR primers

# **DISCUSSION:**

Long noncoding RNAs (lncRNAs) by their number and diversity represent a large field of research and most of their roles are still to be discovered. Many of these lncRNAs have a nuclear localization and among them, some are implicated in regulatory pathways of gene expression through transcriptional or posttranscriptional mechanism. One of the current challenges in this field is to understand the relevance of those lncRNAs in post-transcriptional processing of RNAs. For this purpose, RNAs targeted by lncRNAs have to be identified. Inspired by previous studies focused on association of lncRNAs with chromatin, we developed a procedure that allows the identification of RNAs associated with a lncRNA. The success of this protocol, named RNA pull-down, is mainly dependent on two crucial steps, namely the design of anti-sense DNA oligonucleotide probes that have to specifically and exclusively hybridize with the lncRNA of interest and the conditions of tissue or cell fixation that have to preserve the integrity of the network between all molecular partners.

Previously published protocols provided procedures to isolate a lncRNA together with its associated chromatin sequences (ChIRP (Chu et al 2011; Chu et al 2012), CHART (Simon et al 2011; Simon 2013)). In those protocols, different strategies were employed

to design the anti-sense DNA biotinylated oligonucleotide probes. In ChIRP procedure, the authors used a pool of DNA biotinylated oligonucleotide probes encompassing the entire length of the lncRNA of interest after exclusion of all redundant and non-specific probes (Chu et al 2011; Chu et al 2012). In CHART protocol, the authors identified the regions of the lncRNA that are more accessible for hybridization and design capture oligonucleotides that target these regions. These regions were selected on the basis of their RNase-H sensitivity. Indeed, using the property of RNAse-H to hydrolyze RNAs at sites of RNA-DNA binding, the oligonucleotides that hybridize to accessible sites in the lncRNA produce RNA-DNA hybrids and lead to enzymatic cleavage of the lncRNA. The authors selected three of these candidate capture oligonucleotides and used them in a cocktail (Simon et al 2011; Simon 2013).

The procedure we used to design the anti-sense DNA biotinylated oligonucleotide probes was close to that used in the CHART protocol, but the hybridization-available regions of the desired lncRNA were not selected on the basis of their RNAse-H sensitivity but according to their low probability of internal base pairing as determined by bioinformatics modeling of lncRNA secondary structure. It must be noticed that different secondary structures will be predicted using different algorithms, and probes to be selected should be those that hybridize to available sequences of the lncRNA in the largest number of secondary structures predicted. In our hand, same results were obtained using either a cocktail of three designed specific probes or single probe individually. This prompted us to use separately two specific probes and to consider positive results those that are common to these two probes. Finally, It is therefore recommended at the beginning of the development of the method, for an optimal result and to be able to assess the specificity of results of the pull-down, to design 3 different antisense oligonucleotide probes and then to compare experimentally their efficiency especially since probe efficiency may be altered by cell lysate preparation. Nevertheless, the procedure of probe design based on bioinformatics modeling of lncRNA secondary structure we used remained less expensive than that based on pools of tiling oligonucleotide probes (Chu et al 2012) and less time consuming than that based on RNAse-H sensitivity (Simon 2013).

A negative control should also be performed using as negative capture oligonucleotide either the sense DNA biotinylated oligonucleotide probes or scrambled oligonucleotide probes, or oligonucleotides directed against an unrelated RNA. Because of the existence of natural antisens transcripts lncRNAs, use of sense oligonucleotide probes could be sometimes inadequate. Irrespective of the oligonucleotide probe selected for the negative control, it is necessary to check by blast that it does not hybridize with a known RNA and to keep in mind that this oligonucleotide can hybridize to a still unannotated lncRNA

Cell lysates used in our RNA pull-down experiments were obtained from 10<sup>6</sup> to 10<sup>7</sup> cells when working with cultured cells, and from 1 to 10 mg when working with tissue. The preparation of cell lysates needs to be adapted according to the tissue or cell type used with two main steps that have to be optimized namely the cross-linking step that allows formation of covalent bonds between the lncRNA and its molecular partners and the sonication step that reduces the viscosity by shredding chromatin.

The aim of the cross-linking step is to ensure that all RNA targets remain closed to the lncRNA by inducing the formation of a network between all molecular partners. A paraformaldehyde treatment step that will form covalent bonds between the lncRNA and its partners allows the network to be reticulated. In CHART protocol, it was suggested if

working with nuclear lncRNA, to perform a first treatment with paraformaldehyde on the whole cell lysate and a second treatment on the isolated nucleic fraction (Simon et al 2011; Simon 2013). In our hands, we observed that this supplementary step decreased the efficiency of our probes probably by reducing the lncRNA accessibility in the cells. Hence, the degree of reticulation by paraformaldehyde has to be adapted taking into account the cell or tissue type used, the localization of the lncRNA of interest, and the efficiency of the designed probes.

While lysing the cells, the chromatin is released in the lysate and increases its viscosity; it is then necessary to shred the chromatin by sonication to increase the fluidity of the samples and hence facilitate the accessibility of the oligonucleotide probes to the lncRNA of interest. However sonication will also shred the RNAs extracted with the lncRNA of interest. It is then important to minimize the sonication time in such a way that whereas it efficiently reduces the viscosity of the lysate, it also allows getting RNA fragments with a length comprised between 200-800 bp. It is to note that the sonication time will be highly dependent upon both the quantity and the type of tissue or cultured cells used.

In conclusion, the procedure described here allows in 2-3 days to capture the RNA targets of a desired lncRNA. Coupled with RT-qPCR, it will allow looking for a specific association and regulation of a mRNA by the desired lncRNA as a candidate approach. For a genome-wide approach, RNA pull down experiments can be analyzed by high-throughput RNA-sequencing allowing the retrieval of all RNAs associated with the desired lncRNA. Whatever the analytical strategy chosen, the RNA pull down procedure should provide new significant knowledge on RNA regulation by lncRNAs.

#### **ACKNOWLEDGMENTS:**

This work was supported by Aix-Marseille University and CNRS and funded by a grant from Pfizer Laboratories.

#### **DISCLOSURES:**

The authors have nothing to disclose.

#### **REFERENCES:**

Adriaens, C., Standaert, L., Barra, J., Latil, M., Verfaillie, A., Kalev, P., Boeckx, B., Wijnhoven, P.W., Radaelli, E., Vermi, W., Leucci, E., Lapouge, G., Beck, B., van den Oord, J., Nakagawa, S., Hirose, T., Sablina, A.A., Lambrechts, D., Aerts, S., Blanpain, C. & Marine, J.C., 2016, p53 induces formation of NEAT1 IncRNA-containing paraspeckles that modulate replication stress response and chemosensitivity, *Nat Med*.

Barry, G., Briggs, J.A., Hwang, D.W., Nayler, S.P., Fortuna, P.R., Jonkhout, N., Dachet, F., Maag, J.L., Mestdagh, P., Singh, E.M., Avesson, L., Kaczorowski, D.C., Ozturk, E., Jones, N.C., Vetter, I., Arriola-Martinez, L., Hu, J., Franco, G.R., Warn, V.M., Gong, A., Dinger, M.E., Rigo, F., Lipovich, L., Morris, M.J., O'Brien, T.J., Lee, D.S., Loeb, J.A., Blackshaw, S., Mattick, J.S. & Wolvetang, E.J., 2017, The long non-coding RNA NEAT1 is responsive to neuronal activity and is associated with hyperexcitability states, *Scientific reports*, 7, p. 40127.

Chen, L.L., DeCerbo, J.N. & Carmichael, G.G., 2008, Alu element-mediated gene silencing, *EMBO J*, 27(12), pp. 1694-705.

Chu, C., Qu, K., Zhong, F.L., Artandi, S.E. & Chang, H.Y., 2011, Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions, *Molecular cell*, 44(4), pp. 667-78.

Chu, C., Quinn, J. & Chang, H.Y., 2012, Chromatin isolation by RNA purification (ChIRP), *J Vis Exp*, 61, p. e3912.

Engreitz, J.M., Sirokman, K., McDonel, P., Shishkin, A.A., Surka, C., Russell, P., Grossman, S.R., Chow, A.Y., Guttman, M. & Lander, E.S., 2014, RNA-RNA Interactions Enable Specific Targeting of Noncoding RNAs to Nascent Pre-mRNAs and Chromatin Sites, *Cell*, 159(1), pp. 188-99.

Fang, J., Qiao, F., Tu, J., Xu, J., Ding, F., Liu, Y., Akuo, B.A., Hu, J. & Shao, S., 2017, High expression of long non-coding RNA NEAT1 indicates poor prognosis of human cancer, *Oncotarget*.

Fox, A.H. & Lamond, A.I., 2010, Paraspeckles, *Cold Spring Harbor perspectives in biology*, 2(7), pp. http://cshperspectives.cshlp.org/content/2/7/a000687.long.

Gruber, A.R., Lorenz, R., Bernhart, S.H., Neuböck, R. & Hofacker, I.L., 2008, The Vienna RNA websuite, *Nucleic acids research*, 36(Web Server issue), pp. W70-4.

Nakagawa, S., Ip, J.Y., Shioi, G., Tripathi, V., Zong, X., Hirose, T. & Prasanth, K.V., 2012, Malat1 is not an essential component of nuclear speckles in mice, *RNA*.

Reuter, J.S. & Mathews, D.H., 2010, RNAstructure: software for RNA secondary structure prediction and analysis, *BMC bioinformatics*, 11, p. 129.

Riva, P., Ratti, A. & Venturin, M., 2016, The long non-coding RNAs in neurodegenerative diseases: novel mechanisms of pathogenesis, *Current Alzheimer research*.

Simon, M.D., 2013, Capture hybridization analysis of RNA targets (CHART), *Curr Protoc Mol Biol*, Chapter 21, pp. Unit 21.25..

Simon, M.D., Wang, C.I., Kharchenko, P.V., West, J.A., Chapman, B.A., Alekseyenko, A.A., Borowsky, M.L., Kuroda, M.I. & Kingston, R.E., 2011, The genomic binding sites of a noncoding RNA, *Proceedings of the National Academy of Sciences of the United States of* 

America, 108(51), pp. 20497-502.

Sun, X., Haider Ali, M.S.S. & Moran, M., 2017, The role of interactions of long non-coding RNAs and heterogeneous nuclear ribonucleoproteins in regulating cellular functions, *Biochem J*, 474(17), pp. 2925-35.

Tollervey, J.R., Curk, T., Rogelj, B., Briese, M., Cereda, M., Kayikci, M., König, J., Hortobágyi, T., Nishimura, A.L., Zupunski, V., Patani, R., Chandran, S., Rot, G., Zupan, B., Shaw, C.E. & Ule, J., 2011, Characterizing the RNA targets and position-dependent splicing regulation by TDP-43, *Nat Neurosci*, 14(4), pp. 452-8.

Torres, M., Becquet, D., Blanchard, M.P., Guillen, S., Boyer, B., Moreno, M., Franc, J.L. & François-Bellan, A.M., 2016, Circadian RNA expression elicited by 3'-UTR IRAlu-paraspeckle associated elements, *eLife*, 5.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L. & Pachter, L., 2012, Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks, *Nature protocols*, 7(3), pp. 562-78.

Tripathi, V., Ellis, J.D., Shen, Z., Song, D.Y., Pan, Q., Watt, A.T., Freier, S.M., Bennett, C.F., Sharma, A., Bubulya, P.A., Blencowe, B.J., Prasanth, S.G. & Prasanth, K.V., 2010, The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation, *Molecular cell*, 39(6), pp. 925-38.

Zhang, B., Arun, G., Mao, Y.S., Lazar, Z., Hung, G., Bhattacharjee, G., Xiao, X., Booth, C.J., Wu, J., Zhang, C. & Spector, D.L., 2012, The lncRNA Malat1 is dispensable for mouse development but its transcription plays a cis-regulatory role in the adult, *Cell reports*, 2(1), pp. 111-23.



Figure1





#### Neat1 RNA pull-down



Figure 3



Neat1 RNA pull-down



Figure 4

Malat1 RNA pull-down



Figure 5

# Circadian RNA expression elicited by 3'-UTR IRAlu-paraspeckle associated elements

Manon Torres, Denis Becquet, Séverine Guillen, Bénédicte Boyer, Mathias Moreno, Marie-Pierre Blanchard1, Jean-Louis Franc and Anne-Marie François-Bellan eLife 2016;5:e14837 DOI: 10.7554/eLife.14837

# Paraspeckles as rhythmic nuclear mRNA anchorages responsible for circadian gene expression

Manon Torres, Denis Becquet, Séverine Guillen, Bénédicte Boyer, Mathias Moreno, Marie-Pierre Blanchard1, Jean-Louis Franc and Anne-Marie François-Bellan Nucleus, 2017, Vol. 8, N°. 3, 249-254 DOI : 10.1080/19491034.2016.1277304

#### Introduction

La rythmicité circadienne de la plupart des fonctions cellulaires, physiologiques et comportementales repose sur l'expression rythmique de nombreux gènes dans tout l'organisme. Le mécanisme moléculaire à la base du fonctionnement des oscillateurs circadiens implique des boucles de rétrocontrôle de la transcription, et ce mécanisme transcriptionnel qui permet le contrôle de la rythmicité d'expression de nombreux gènes cibles des oscillateurs a été très étudié. Cependant plusieurs études récentes ont permis de montrer l'importance majeure des mécanismes post-transcriptionnels dans le contrôle de la rythmicité d'expression des gènes cibles (Menet et al 2012; Koike et al 2012).

Au cours de notre étude nous nous sommes intéressés à l'implication possible d'un mécanisme post-transcriptionnel basé sur la rétention nucléaire d'ARNm dans le contrôle de la rythmicité circadienne d'expression de gènes. Ce mécanisme implique des structures nucléaires appelées les paraspeckles. Les paraspeckles sont des corps nucléaires formés autour du lncRNA Neat1 auquel se fixent plusieurs protéines présentant des sites de liaison aux ARNs (Fox et al 2002; Bond, & Fox 2009). Du fait de leur découverte récente, leurs fonctions physiologiques ainsi que le mécanisme de leur fonctionnement et leurs implications possibles en pathologie sont encore mal déterminés. On sait néanmoins qu'ils ont la capacité de retenir dans le noyau certains ARNs, empêchant ainsi leur export cytoplasmique et donc leur traduction en protéine (Prasanth et al 2005; Sasaki et al 2009; Chen, & Carmichael 2009) . La rétention de ces ARNs dépendrait de structures particulières présentes dans leur région 3' non codante (3'UTR). Ces structures appelées IRAlu correspondent à deux séquences Alu qui sont des séquences spécifiques de l'homme et des primates, appartenant à la famille des SINE que l'on trouve en grand nombre dans la région 3'UTR, qui sont inversées et peuvent ainsi s'apparier pour former une boucle d'ARN double brin. Cette boucle d'ARN double brin est reconnue par des enzymes d'édition et les motifs IRAlus dans lesquels l'Adénosine est éditée en Inosine, sont reconnus et retenus dans le noyau par les paraspeckles (Prasanth et al 2005; Chen et al 2008).

Des travaux antérieurs réalisés au sein de l'équipe ont montré que deux des protéines principales des paraspeckles, NONO et SFPQ, sont exprimées avec une rythmicité circadienne dans la lignée hypophysaire GH4C1 (Guillaumond et al 2011). Nous avons recherché si dans les cellules GH4C1, l'ensemble des éléments des paraspeckles était rythmique et si cette rythmicité pouvait concourir à une rétention nucléaire rythmique des ARNs reconnus par ces structures.

#### **Résultats**

Dans cette étude nous avons montré l'existence de paraspeckles dans la lignée GH4C1 et nous en avons caractérisé la structure par une étude de microscopie en super résolution STORM. Nous avons montré que les éléments principaux, Neat1 et les protéines NONO, SFPQ, PSPC1 et RBM14 sont exprimés de manière rythmique et leur association suit aussi une rythmicité circadienne. De ce fait le nombre de paraspeckles au sein des cellules, évalué en microscopie confocale après Hybridation in situ de fluorescence (FISH) de Neat1, varie de manière circadienne.

Une séquence IRAlu a été insérée avant le site de polyadénylation en 3'-UTR de l'ARNm de la Green Fluorescent Protein (GFP). Après transfection de cette construction dans des cellules GH4C1 et sélection de lignées exprimant de manière stable cette construction, nous avons montré que l'ajout du motif IRAlu induit l'association de l'ARNm de l'egfp avec les différents éléments des paraspeckles. Cette association est corrélée à un rythme circadien de la rétention nucléaire de l'ARNm egfp associé à un rythme circadien de l'expression cytoplasmique de la protéine EGFP détecté par vidéomicroscopie en temps réel. Lorsque les paraspeckles sont détruits après transfection des cellules par des siRNA ou des oligonucléotides antisens dirigés contre Neat1, ces rythmes disparaissent.

Par l'utilisation de la construction IRAlu-egfp nous avons donc montré que les paraspeckles peuvent réguler au niveau post-transcriptionnel l'expression circadienne d'ARNs en les retenant dans le noyau.

Notre étude a ensuite visé à déterminer la contribution de Neat1 via les paraspeckles au transcriptome rythmique hypophysaire. En couplant la technique du RNA pulldown avec du séquençage haut débit des ARNs, il a été évalué qu'environ 28% des ARNs exprimés dans les cellules GH4C1 se trouvent associés à Neat1. Plusieurs ARNs sélectionnés au sein de cette liste ont été testés et pour tous la liaison à Neat1 est associée à une rétention nucléaire qui suit un profil circadien. Le Knock-down de Neat1 induit pour chacun de ces ARNs une perte de la rythmicité de la rétention nucléaire, confirmant ainsi que la rythmicité circadienne d'expression de plusieurs ARNs endogènes est contrôlée au niveau post-transcriptionnel par les paraspeckles.

# Conclusion

Au cours de cette étude, nous avons pu montrer la présence et la rythmicité de formation des paraspeckles dans un modèle de cellules hypophysaires. Du fait de leur rythmicité, les paraspeckles vont induire la rétention nucléaire circadienne d'un gène rapporteur egfp présentant le motif IRAlu dans sa région 3'UTR ainsi que l'expression rythmique de la protéine EGFP correspondante. Dans le modèle cellulaire GH4C1, la rythmicité d'expression de plusieurs ARNs endogènes dépend de ce mécanisme et est abolie lorsque la formation des paraspeckles est inhibée. Grâce à la méthode de RNA pull-down couplée à du séquençage haut débit des ARNs, il a pu être montré que plus d'un quart des ARNs exprimés dans la lignée GH4C1 se lient à Neat1. Ces résultats suggèrent donc que Neat1 et les paraspeckles pourraient jouer un rôle majeur dans la régulation de la rythmicité circadienne du transcriptome hypophysaire.



# Circadian RNA expression elicited by 3'-UTR IRAlu-paraspeckle associated elements

Manon Torres, Denis Becquet, Marie-Pierre Blanchard, Séverine Guillen, Bénédicte Boyer, Mathias Moreno, Jean-Louis Franc, Anne-Marie François-Bellan\*

Faculté de Médecine Nord, Aix Marseille Université, CNRS, CRN2M-UMR7286, Marseille, France

**Abstract** Paraspeckles are nuclear bodies form around the long non-coding RNA, Neat1, and RNA-binding proteins. While their role is not fully understood, they are believed to control gene expression at a post-transcriptional level by means of the nuclear retention of mRNA containing in their 3'-UTR inverted repeats of Alu sequences (IRAlu). In this study, we found that, in pituitary cells, all components of paraspeckles including four major proteins and Neat1 displayed a circadian expression pattern. Furthermore the insertion of IRAlu at the 3'-UTR of the EGFP cDNA led to a rhythmic circadian nuclear retention of the egfp mRNA that was lost when paraspeckles were disrupted whereas insertion of a single antisense Alu had only a weak effect. Using real-time video-microscopy, these IRAlu were further shown to drive a circadian expression of EGFP protein. This study shows that paraspeckles, thanks to their circadian expression, control circadian gene expression at a post-transcriptional level.

DOI: 10.7554/eLife.14837.001

# Introduction

The circadian clock orchestrates daily rhythms in metabolism, physiology and behavior that allow organisms to anticipate regular changes in their environment, increasing their adaptation (Asher and Schibler, 2011). Circadian rhythms are underpinned by daily rhythms of gene expression. The transcriptional component of these rhythms is well understood (Zhang and Kay, 2010). The circadian variation in abundance of the positive (Clock and Bmal1) and the negative (Per1, Per2 and Cry1, Cry2) components of these loops drive the circadian transcription of both direct targets genomewide and a cascade of circadian output transcription factors, which together mediate the circadian transcriptional profile of a cell type or tissue (Asher and Schibler, 2011; Rey et al., 2011). Though the core circadian system has concentrated on transcriptional control, it has been apparent that substantial regulation is achieved after transcription so that post-transcriptional controls are emerging as crucial modulators of circadian clocks (Lim and Allada, 2013; Wang et al., 2013; Menet et al., 2012; Koike et al., 2012; Hurley et al., 2014). While in eukaryotes, approximately 1%-10% genes are subjected to circadian control directly or indirectly only ~1/5 of the mRNAs that display rhythmic expression are directly driven by transcription, which suggests that post-transcriptional mechanisms including RNA splicing, polyadenylation, mRNA stability, mRNA cytoplasmic export and RNAs nuclear retention are essential layers for generation of gene expression rhythmicity (Partch et al., 2014; Menet et al., 2012; Koike et al., 2012; Hurley et al., 2014).

Paraspeckles are recently identified nuclear bodies that have been shown to retain RNAs in the nucleus. Paraspeckles contain proteins PSPC1, RBM14, NONO, and SFPQ (*Prasanth et al., 2005*) and are usually detected as a variable number of discrete dots found in close proximity to nuclear speckles (*Bond and Fox, 2009*). One long noncoding RNA, nuclear-enriched abundant transcript

\*For correspondence: annemarie.francois@univ-amu.fr

**Competing interests:** The authors declare that no competing interests exist.

Funding: See page 20

Received: 29 January 2016 Accepted: 20 July 2016 Published: 21 July 2016

**Reviewing editor:** Robert H Singer, Albert Einstein College of Medicine, United States

© Copyright Torres et al. This article is distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use and redistribution provided that the original author and source are

credited.

CC

**eLife digest** Many biological features of animals, including body temperature and hormone levels, follow daily rhythms that repeat every 24 hours. These so-called circadian rhythms are driven by an internal body clock and are essential for the organism to adapt to the daily cycle of light and dark. Circadian rhythms also take place inside individual cells – for example, the amount of a given protein in a cell often rises and falls over each 24-hour period.

To generate these daily fluctuations, the processes used to make proteins based on the instructions encoded within a gene must be carefully controlled. Genes are first copied or 'transcribed' into intermediate molecules called messenger RNAs (mRNAs). These mRNA molecules must then travel out of the cell's nucleus before they can be de-coded to produce proteins. This means that daily fluctuations in mRNA and protein levels could occur because the rate at which the DNA is transcribed fluctuates or because controlling the steps that occur after transcription. However it is not clear how much these post-transcriptional steps contribute to circadian rhythms inside cells.

Recently, structures called paraspeckles were seen inside the nucleus. These structures are made from a long RNA molecule that does not code for a protein, and a number of proteins that can bind mRNA molecules. Paraspeckles are thought to prevent certain mRNAs from leaving the nucleus and therefore stop them from being decoded to make proteins. Torres et al. have now investigated whether paraspeckles may play a role in circadian rhythms.

Torres et al. looked at the long non-coding RNA and several proteins that are known to be components of paraspeckles in cells taken from the pituitary glands of rats using a variety of techniques. These cells were chosen because they were known to have a working circadian clock. The analysis showed that the levels of these components, as well as the number of paraspeckles within the nucleus, changed over the course of a daily cycle.

Torres et al. then confirmed that mRNAs containing a sequence that is known to recruit mRNAs to paraspeckes (the IRAlu sequence) could be also retained in the nucleus or released with a circadian rhythm. This pattern was lost when the paraspeckles were disrupted.

These findings suggest that daily fluctuations in protein levels can be post-transcriptionally controlled by paraspeckles rhythmically retaining mRNAs in the nucleus. Future studies could explore whether it may be possible to control circadian rhythms by targeting the paraspeckles, which could help to improve conditions where the internal body clock goes wrong. DOI: 10.7554/eLife.14837.002

one (Neat1), exclusively localized to paraspeckles serves as a structural component (*Hutchinson et al., 2007; Chen and Carmichael, 2009; Clemson et al., 2009; Sasaki and Hirose, 2009; Sunwoo et al., 2009*). The locus generates short and long transcripts from the same promoter, which have previously been identified as MEN $\epsilon$  (Neat1-1) and MEN $\beta$  (Neat1-2), respectively (*Guru et al., 1997; Sasaki and Hirose, 2009*). Because specific depletion of Neat1-2 leads to disruption of paraspeckles (*Sasaki and Hirose, 2009*) Neat1-1 alone cannot induce paraspeckle formation.

Paraspeckles have been shown to retain in the nucleus RNAs containing duplex structures (*Chen and Carmichael, 2008*). This is the case for the mouse cationic amino acid transporter 2 (Cat2) transcribed nuclear RNA, Ctn-RNA, an alternatively spliced form of the Cat2 mRNA, which contains a dsRNA structure resulting from inverted short inter-spersed nuclear elements (SINEs) in its 3'-UTR (*Prasanth et al., 2005*). In human cells, hundreds of genes contain inverted repeated SINEs (mainly Alu elements) in their 3'-UTRs. Alu elements are unique to primates and account for almost all of the human SINEs and >10% of the genome. Their abundance leads to the frequent occurrence of inverted repeat structures (inverted repeated Alu elements [IRAlus]) in gene regions (*Chen et al., 2008*). It has been reported previously that mRNAs containing IRAlus in their 3'-UTRs like Nicolin 1 (NICN1) or Lin 28 are retained in the nucleus in paraspeckles (*Chen et al., 2008*). Therefore, this nuclear retention pathway of IRAlus in 3'-UTRs of genes provides an additional layer of gene regulation by sequestering mature mRNAs within the nucleus.

We recently reported that two protein components of paraspeckles, namely NONO and SFPQ, display a circadian expression pattern in primary cultures of pituitary cells as well as in a rat pituitary cell line, the GH4C1 cells (*Becquet et al., 2014; Guillaumond et al., 2011*). We used this cell line to determine whether one of the posttranscriptional mechanisms allowing circadian gene expression in pituitary cells could involve the circadian nuclear mRNA retention by paraspeckle bodies. To this end, we first characterized the presence of paraspeckles and we showed that these nuclear bodies were rhythmically expressed in the rat GH4C1 pituitary cell line. We then made a series of EGFP-fused IRAlu or Alu constructs and transfected them into GH4C1 cells to investigate the EGFP expression and the fates of the egfp-IRAlu or egfp-Alu RNA. We showed that an IRAlu element in the 3'-UTR of the egfp mRNA strongly repressed EGFP expression. Further, this reduction was accompanied by significant nuclear retention of the mRNAs, likely by paraspeckle bodies. We showed also that insertion of IRAlus in the 3'-UTR of EGFP reporter gene allowed rhythmic nuclear egfp-IRAlu RNA retention and rhythmic EGFP protein expression. Finally, this rhythmic nuclear egfp-IRAlu RNA retention as well as the rhythmic nuclear retention of some known cycling transcripts that were shown here associated with paraspeckles was lost when paraspeckles were disrupted.

# **Results**

# Characterization of paraspeckle nuclear bodies in pituitary GH4C1 cells

Visualization of paraspeckle protein components by immunofluorescence and biochemical evidence for their association with the long noncoding RNA Neat1

The presence of paraspeckles in GH4C1 cells was anatomically evidenced by confocal microscopy on the basis of the overlap of their protein components using antibodies directed against SFPQ, NONO, PSPC1 and RBM14. SFPQ (green Figure 1A,C), NONO (green Figure 1B,D), PSPC1 (red Figure 1A,B) and RBM14 (red Figure 1C,D) staining appeared as punctate clusters located exclusively within the boundaries of the nucleus as delimitated by Hoechst labeling (Figure 1A,B,C,D, 4th column). When SFPQ or NONO staining was merged with PSPC1 (Figure 1A,B, 3rd and 4th columns) on the one hand or RBM14 (Figure 1C,D, 3rd and 4th columns) on the other hand, some punctate clusters were found to overlap with each other in subnuclear structures reminiscent of paraspeckles. The protein components of paraspeckles were further shown to be associated with the long noncoding RNA Neat1IncRNA Neat1, using RNA-immunoprecipitation (RIP) experiments with antibodies directed against NONO, SFPQ, RBM14 or PSPC1. Primers used in RT-qPCR detected both long (Neat1-2) and short (Neat1-1) isoforms of Neat1 RNA (Figure 2-source data 1). A 20- to 80-fold enrichment in Neat1 RNA was obtained with these antibodies as compared to an irrelevant non-specific antibody (Figure 2A). These results showed that the four major protein components of paraspeckles bound the structural IncRNA Neat1, confirming the presence of paraspeckle nuclear bodies in GH4C1 cells.

# Visualization of paraspeckle nuclear bodies by fluorescence in situ hybridization (FISH) of long noncoding RNA Neat1

Paraspeckle nuclear bodies were then visualized using Neat1 RNA staining by FISH. Under confocal laser scanning microscope, Neat1 RNA FISH staining appeared as regular punctates within the boundaries of the nucleus (*Figure 2B*). Number of foci was low and variable from one cell to the other. To gain resolution in the intranuclear spatial arrangement of Neat1 RNA, we combined RNA FISH with Stochastic Optical Reconstruction Microscopy (STORM). Under conventional fluorescence microscopy, resolution of paraspeckle bodies was very low as can be seen in the left part and in upper panel of the right part (red staining) in *Figure 2C*. The increase in resolution from conventional to super-resolution microscopy was apparent in the upper panel of the right part of *Figure 2C* in which the high-resolution image obtained after STORM analysis (white staining) was superimposed to the conventional fluorescence image (red staining). Due to the poor resolution, the size of the paraspeckles was imprecise and could not be measured under conventional fluorescence microscope (red staining). By contrast, dimensions could be precisely measured under STORM (bottom panel of the right part in *Figure 2C*). In addition to the precision in the real size of paraspeckle bodies,



**Figure 1.** Localization by confocal microscopy of paraspeckle proteins in GH4C1 pituitary cells. Cells are grown on coverslips and labeled for immunofluorescence with antibody to SFPQ (green A, C), NONO (green B, D) PSPC1 (red A, B) or RBM14 (red C, D). SFPQ or NONO staining is merged with PSPC1 (A, B 3th column) or RBM14 (C, D 3th column). Nuclear staining by Hoechst for the same samples is added in 4th column. Arrows indicate punctate clusters in which two paraspeckle proteins overlap. Scale bars: 5 μm. DOI: 10.7554/eLife.14837.003

super-resolution analysis allowed also to show that Neat1 labeling of paraspeckle bodies was not round in shape, being merely elliptical. Indeed using an imaging software (NIS-Elements, Nikon France S.A, Champigny sur Marne, France) we showed that the averaged ratio of Height to Width (n=10) was >1 (1.53  $\pm$  0.1) and the mean surface area was 24,960  $\pm$  4831 nm<sup>2</sup> (*Figure 2C*).

#### Rhythmic expression of paraspeckle components in GH4C1 cells

#### Rhythmic expression of protein components

To determine whether paraspeckle nuclear bodies displayed a circadian expression pattern we looked for rhythmic expression of their protein components as determined by Western blot analysis in nuclear protein extracts. We previously reported that the two proteins NONO and SFPQ followed a rhythmic expression pattern in GH4C1 cells (*Guillaumond et al., 2011*). It holds also true for two other paraspeckle-associated proteins, RBM14 and PSPC1, as reported in *Figure 3A*. Indeed, RBM14 and PSPC1 proteins displayed a rhythmic pattern in GH4C1 cells over the T2-T30 time period that could be fitted with a non-linear cosinor fit in which the period value (2pi/Frequency) was constrained to the circadian period value 24 hr (equation values given in *Figure 3—source data* 1). It may be noticed that the rhythmic expression pattern of RBM14 and PSPC1 proteins (*Figure 3A*) matched quite correctly with that of NONO and SFPQ proteins we previously reported (see Figure 6A in *Guillaumond et al., 2011*).


Figure 2. Neat1 RNA in paraspeckle nuclear bodies: association with paraspeckle proteins and visualization by FISH. (A) Association of paraspeckle proteins with Neat1 RNA RNA Immuno-Precipitation (RIP) experiments (n=4 for each antibody) using antibodies directed against NONO, SFPQ, RBM14 or PSPC1 show the enrichment in Neat1 RNA obtained as compared to an irrelevant non-specific antibody (see Figure 2-source data 1 for primer sequences). \*\*p<0.01 vs non-specific antibody. B. Visualization of Neat1 RNA by Fluorescence in situ Hybridization and confocal laser scanning microscope Left Panel: RNA-FISH shows the distribution of Neat1 RNA in a few distinct foci (arrows). The round aspect of the foci under confocal laser scanning microscope is shown in the insert in which assigned foci is enlarged. Scale bars: 1 µm Right Panel: The extent of the nucleus is shown with Hoechst staining. Foci containing Neat1 RNA localize within the nucleus sometimes in the close vicinity of nucleus boundaries. Scale bars: 5 µm. (C) Visualization of Neat1 RNA by Fluorescence in situ Hybridization and super resolution Left Panel: Conventional fluorescence microscopy of Neat1 RNA-FISH. The nucleus is outlined with hand-drawn dashed lines to indicate the nuclear periphery. Scale bars: 5 µm. Right Upper Panel: Enlargement of the foci assigned in left panel allows to show the poor resolution of paraspeckle under conventional fluorescence microscopy (in red). In white is the superimposed high-resolution image obtained after STORM analysis. Note that the size of paraspeckle after STORM analysis is strongly reduced. Scale bars: 0.5 µm. Right Bottom Panel: Enlargement of the STORM analysis shown in the upper panel with measurements of width (Wstorm=0.14 µm) and height (hstorm=0.17  $\mu$ m). Note the elliptical shape of the foci analyzed. Scale bars: 0.1  $\mu$ m. DOI: 10.7554/eLife.14837.004

The following source data is available for figure 2:

**Source data 1.** Sequences of qPCR primers and oligonucleotides. DOI: 10.7554/eLife.14837.005

# Rhythmic expression of Neat1 RNA

We further looked for a circadian expression pattern of the structural element of paraspeckles, namely the lncRNA Neat1. As shown in *Figure 3B*, Neat1 RNA levels (Neat1-1 + Neat1-2) displayed a rhythmic pattern in GH4C1 cells over the T2-T38 time period (cosinor fit values given in *Figure 3— source data 1*). Neat1 RNA levels (Neat1-1 + Neat1-2) also displayed a circadian expression pattern in several mouse tissues including the pituitary gland but also other peripheral oscillators such as the spleen or the adrenal gland (*Figure 3—figure supplement 1*). This was also the case in the central clock, namely the suprachiasmatic nuclei (*Figure 3—figure supplement 1*).

# Rhythmic binding of paraspeckle proteins on Neat1 RNA

We used RIP experiments to examine whether binding of NONO, SFPQ, RBM14 and PSPC1 on Neat1 RNA in GH4C1 cells was rhythmic. As shown in *Figure 3C*, binding of the four paraspeckle-associated proteins on Neat1 RNA displayed a rhythmic pattern (cosinor fit values given in *Figure 3—source data 2*). Maximum binding on Neat1 RNA was reached between 6 and 10 hr after the medium change for the four proteins (*Figure 3C*).



Figure 3. Rhythmic expression and association of paraspeckle components. (A) Rhythmic expression of two paraspeckle proteins in pituitary GH4C1 cells The expression of PSPC1 and RBM14 is determined by Western Blot analysis over a 30 hr time period. Each data point (mean ± SD of three independent samples) represents the ratio of the depicted proteins to ATF2 and is expressed relative to the value obtained at ZT 30. Experimental values can be adequately fitted ( $R^2$ >0.55) with a non-linear cosinor equation in which the period value is set to 24 hr (see also Figure 3-source data 1). (B) Rhythmic expression of the long noncoding Neat1 RNA. The expression of the IncRNA Neat1 is determined by RT-qPCR over a 40 hr time period. Primers used to allow the detection of both Neat1-1 and Neat1-2. Experimental values (n=4) expressed as a percent of the initial value obtained at ZT 0 can be adequately fitted ( $R^2$ >0.55) with a non-linear cosinor equation in which the period value is set to 24 hr (see also Figure 3-source data 1). (C) Rhythmic association of paraspeckle proteins with Neat1 RNA RNA Immuno-Precipitation (RIP) experiments (n=4 for each antibody) are performed over a 30h time period. At each time point, the levels of Neat1 RNA determined after immuno-precipitation by the antibodies directed against NONO, SFPQ, RBM14 and PSPC1 were normalized relative to Neat1 RNA input levels and expressed as a percent of the value obtained at T0. Experimental values can be adequately fitted (R<sup>2</sup>>0.55) with a non-linear cosinor equation in which the period value is set to 24 hr (see also Figure 3-source data 2). (D-E) Rhythmic fluctuations of paraspeckle number Cells were arrested at four different times after the medium change and processed for FISH of Neat1 RNA. At each time point, 20 to 35 images from four wells of 100 000 cells obtained in two different experiments were acquired under a confocal microscope with a 40X objective. At each time point, the total number of paraspeckles per well and the mean number of paraspeckles per cell were calculated. \*\*p<0.001 \*\*\*\*p<0.0001. Figure 3 continued on next page





# Rhythmic expression of the long noncoding Neat1 RNA in diverse circadian oscillators.

Daily pattern of expression of the long noncoding Neat1 RNA in different tissue of mice maintained in a 12h/12h light-dark cycle and sacrificed at different Zeitgeber Time (ZT). ZT0 corresponds to time of light on. Total RNA was prepared and used for cDNA synthesis followed by RT-qPCR (see Figure 2—source data 1 for the list of primers). Primers used to allow the detection of both Neat1-1 and Neat1-2. Neat1 RNA accumulation was normalized relative to 36B4 mRNA levels. White and black bars designate the light and dark phase, respectively. Experimental values expressed as a percent of the initial value obtained at ZT 0 can be adequately fitted (R2>0.55) with a non-linear sine wave equation in which the period value is set to 24 hr.

#### Figure 3 continued

#### DOI: 10.7554/eLife.14837.006

The following source data and figure supplement are available for figure 3:

**Source data 1.** Cosinor analysis of the rhythmic expression pattern of paraspeckle components in GH4C1 cells. DOI: 10.7554/eLife.14837.007

**Source data 2.** Cosinor analysis of the rhythmic binding of the four paraspeckle-associated proteins on Neat1 RNA in GH4C1 cells.

#### DOI: 10.7554/eLife.14837.008

**Figure supplement 1.** Rhythmic expression of the long noncoding Neat1 RNA in diverse circadian oscillators. DOI: 10.7554/eLife.14837.009

# Rhythmic number of paraspeckles

In cells arrested at four different times after the medium change, the total number of paraspeckles per well was shown to fluctuate, being significantly lower 15 hr after the medium change i.e. at a time when Neat1 RNA levels were around the lowest levels (*Figure 3D*,  $F_{3,103}$ =9.531 p<0.0001). Furthermore at this time point, the mean number of paraspeckles per cell reached a minimum value (*Figure 3D*,  $F_{3447}$ = 5.456, p<0.001). Taken together, these results showed that the rhythm of Neat1 and its associated proteins reported above translate into a rhythm in the number of paraspeckles inside the cells.

### Influence of IRAlu elements inserted in 3'-UTR EGFP mRNA

IRAlu elements reduced EGFP protein expression

Paraspeckles have been shown to retain in the nucleus RNAs containing duplex structures from inverted repeats of the conserved Alu sequences (IRAlu elements) within their 3'-UTR (*Chen and Carmichael, 2008*). This has been shown to be the case for Nicolin 1 (NICN1) gene (*Chen and Carmichael, 2008*). We utilized EGFP expression reporter system to investigate the effects of IRAlu from the 3'-UTR of NICN1 gene. The single antisense Alu, or the IRAlu elements cloned from the 3'-UTR of NICN1 were inserted each between the EGFP cDNA 3'-UTR region and the SV40 polyadenylation signal of the expression vector pEGFP-C1 to generate constructs that were then stably transfected into GH4C1 cells. In agreement with previous results by Chen et al (*Chen and Carmichael, 2008*), the IRAlu elements derived from NICN1 significantly reduced EGFP expression when compared with the Alu element (*Figure 4A–C*). This is here evidenced both by the significant reduction in the number of fluorescent cells (*Figure 4B*) and by the significant decrease in relative EGFP levels measured by western blotting in IRAlu-egfp cell line compared to both Alu–egfp cell line (*Figure 4C*). By contrast, we have not observed any difference in gene expression between the parent plasmid pEGFP and derivatives that contained a single Alu element (data not shown).

#### IRAlu element induced egfp mRNA nuclear retention

The reduced EGFP protein expression in IRAlu-egfp cell line suggested that IRAlu can induce a stronger nuclear retention of egfp mRNA than Alu did, as previously shown by Chen et al (*Chen and Carmichael, 2008*). Given that egfp mRNA cytoplasmic localization correlates strongly with EGFP expression, we confirmed in our cell lines that after fractionating cytoplasmic and nuclear RNAs, IRAlu-containing egfp mRNA appeared to be preferentially retained in the nucleus in comparison with Alu-containing egfp mRNA. As shown in *Figure 4D*, the IRAlu from Nicn1 caused a more than two-fold greater nuclear retention of the egfp mRNA when compared with the corresponding Alu element. On the other hand, there is no significant difference in nuclear/cytoplasmic distribution of egfp-Alu mRNA compared with control (data not shown), consistent with our finding that a single Alu element in the 3'-UTR does not affect EGFP gene expression. It then appears that nuclear retention of IRAlu-containing egfp mRNA correlates with silencing of EGFP protein expression (*Figure 4A–C*).



**Figure 4.** Influence of IRAlu elements inserted in 3'-UTR egfp mRNA. (A–C) Decrease in EGFP expression by insertion of IRAlus in the 3' -UTR of egfp mRNA. IRAlus and Alu were PCR-amplified from the 3'-UTR of Nicn1 and then inserted separately into the 3'-UTR of egfp mRNA. GH4C1 cell lines, Alu-Figure 4 continued on next page



#### Figure 4 continued

egfp and IRAlu-egfp cell lines were established by transfection of the indicated plasmids. (A) Representative example of fluorescence and corresponding bright field pictures taken 48 hr after platting of each cell line. Scale bars equal 50  $\mu$ m. (B) Quantitative analysis of the percent of fluorescent cells in each cell line. Data are means ± SEM of 18 measures performed in 3 independent experiments. (C) Quantification of relative levels of eGFP investigated by western blotting with anti-GFP antibody in total protein extracts from the two cell lines. Tubulin was used as the loading control. Data are mean ± SEM of values obtained from three experiments in IRAlu-egfp cell line and are expressed as a percent of the corresponding value obtained in Alu-egfp cell line. (D) Nuclear and cytoplasmic egfp mRNA were quantified by qPCR in each cell line and normalized to the relative amount of gapdh mRNA (n=8 for each cell line). Ratio of nuclear versus cytoplasmic egfp mRNA levels are compared between IRAlu-egfp and Alu-egfp cell lines. (E) Enrichment in IncRNA Neat1 after RNA Immuno-Precipitation (RIP) with an antibody directed against PSPC1 relative to a non-specific oligonucleotide (two right panels). The relative enrichment in IncRNA Neat1 obtained after either RIP (n=3 for each cell line) or RNA pull-down (n=6 for each cell line) is not statistically different in Alu-egfp versus IRAlu-egfp cell lines F. Enrichment in IRAlu-egfp versus Alu-egfp mRNA obtained after either RIP (n=3 for each cell line) or RNA pull-down (n=6 for each cell line) is statistically different to a non-specific oligonucleotide (two right panels). The relative to an irrelevant antibody (left panel) or after RNA pull-down (n=6 for each cell line) is statistically different in Alu-egfp versus IRAlu-egfp cell lines F. Enrichment in egfp mRNA after RNA Immuno-Precipitation (RIP) with an antibody directed against PSPC1 relative to an irrelevant antibody (left panel) or after RNA pull-down with two different specific biotinylated oligonucleotides (S o

DOI: 10.7554/eLife.14837.010

The following figure supplements are available for figure 4:

**Figure supplement 1.** Secondary structure of the first 2500 nucleotides of the Neat1 RNA as predicted by MaxExpect software. DOI: 10.7554/eLife.14837.011

**Figure supplement 2.** Enrichment of Neat1 RNA relative to input after RNA pull-down. DOI: 10.7554/eLife.14837.012

### IRAlu element associated with paraspeckle components IRAlu element associated with PSPC1 protein

We also asked whether IRAlu-containing egfp mRNA is associated with paraspeckle protein PSPC1. To answer this question, we performed RIP experiments on Alu-egfp and IRAlu-egfp cell lines using PSPC1 antibody compared to irrelevant antibody. While the enrichment in IncRNA Neat1 obtained after use of PSPC1 antibody was comparable in the two cell lines (*Figure 4E*), enrichment in egfp mRNA was significant only in IRAlu-egfp cell line (*Figure 4F*), attesting that PSPC1 protein was associated to IRAlu-egfp and not to Alu-egfp mRNA.

#### IRAlu element associated with IncRNA Neat1

We next asked whether IRAlu-containing eqfp mRNA is associated with endogenous IncRNA Neat1. To answer this guestion, we adapted a pull-down Neat1 technology from two published approaches (RIA (RNA-interactome analysis) and CHART (Capture Hybridization Analysis of RNA Targets) (Kretz et al., 2013; Simon, 2013; West et al., 2014). This technology was based on an affinity purification of Neat1 RNA together with its protein partners and mRNA targets by using oligonucleotides that are complementary to Neat1 RNA sequences. In our adapted technology, anti-sense oligonucleotides were designed in stretches of Neat1 RNA available for hybridization and not occluded by in silico predicted secondary structure (Figure 4-figure supplement 1). With this strategy, we found that two anti-sense oligonucleotides allowed a 30 to 40-fold enrichment in Neat1 RNA compared to Neat1 RNA input from cross-linked GH4C1 cellular extracts (Figure 4-figure supplement 2). We performed Neat1 RNA pull-down using these two specific biotinylated complementary oligonucleotides that target Neat1 and one biotynylated irrelevant probe in cell lines stably expressing Alu-containing egfp mRNA or IRAlu-containing egfp mRNA. While one of the specific oligonucleotide (S oligo 2) is more efficient than the other (S oligo 1) to pull-down Neat1, the enrichment in Neat1 was not statistically different in the two Alu-egfp and IRAlu-egfp cell lines (Figure 4E). By contrast, the amounts of egfp mRNA retrieved after Neat1 RNA pull-down by the two specific probes compared to the non-specific probe were significantly higher in IRAlu-egfp compared to Alu-egfp cell line (Figure 4F). This result clearly showed that IRAlu-egfp mRNA was preferentially associated with Neat1-containing paraspeckles compared to Alu-egfp mRNA.



# Secondary structure of the first 2500 nucleotides of the Neat1 RNA as predicted by MaxExpect software.

(A) The structure is colored according to a base-pairing probability code (zoom to view base colors). (B) Enlargement of the red box in A. The two anti-sens oligonucleotide probes designed (S oligo 1 and S oligo 2 in red) are positioned along the Neat1 RNA structure.



Neat1 RNA pull-down

# Enrichment of Neat1 RNA relative to input after RNA pull-down.

Enrichment obtained after using the two anti-sens specific (S oligo 1 and S oligo 2) or a non-specific (NS) oligonucleotide probes. \*\*p<0.01 \*\*\*p<0.001vs non-specific oligonucleotide probe.

### IRAlu element induced egfp mRNA circadian nuclear retention

Since we showed that paraspeckles displayed a circadian expression in GH4C1 cells, we asked whether IRAlu-egfp mRNA shown here to be associated with paraspeckles could be rhythmically retained in the nucleus. To answer this question, we fractionated cytoplasmic and nuclear RNAs in IRAlu-egfp and Alu-egfp cells harvested every 4 hr during 44 hr. We showed that in addition to the previously described higher ratio of nuclear versus cytoplasmic egfp mRNA in IRAlu-egfp cell line (*Figure 4D*) we also found high amplitude circadian variations in this ratio in IRAlu-egfp cell line (*Figure 5A*). Whereas it was observed that the nuclear/cytoplasmic ratio of egfp mRNA could also be fitted by a cosinor equation in Alu-egfp cell line, the magnitude of this rhythmic pattern was highly reduced (*Figure 5A*) as compared to IRAlu-egfp cell line (*Figure 5-* source data 1).

# Loss of circadian nuclear egfp-IRAlu RNA retention after paraspeckle disruption

To address the issue whether the presence of paraspeckles was necessary to the rhythmic egfp mRNA nuclear retention in IRAlu-egfp cell line, we employed two strategies to knockdown Neat1 RNA, namely small interfering RNA (siRNA) and antisens oligonucleotides (ASO). Neat1 has been previously shown by others to be effectively knocked down by siRNA (Chen and Carmichael, 2009; Clemson et al., 2009), and we indeed verified by RT-qPCR that Neat1 RNA levels were reduced in IRAlu-egfp cell line to around 60% after treatment with specific siRNA compared to negative control siRNA, a decrease that was comparable to that obtained with Neat1 ASO (Figure 5-figure supplement 1A); more importantly, whereas Neat1 rhythmic expression pattern was not affected by negative control (Figure 5-figure supplement 1B), it was suppressed both by Neat1 siRNA and Neat1 ASO (Figure 5-figure supplement 1C,D). In addition, when IRAlu-egfp cells were transfected either by Neat1 siRNA or Neat1 ASO as compared to negative control, the relative ratio of nuclear versus cytoplasmic eqfp mRNA levels was significantly decreased (Figure 5B). Moreover, whereas the circadian egfp mRNA nuclear retention was not affected after transfection of the cells by the negative control (not shown), it was abolished after both Neat1 siRNA and Neat1 ASO transfection (Figure 5C). These loss of function experiments showed that the IRAlu-egfp reporter does not cycle when paraspeckles are disrupted.

#### IRAlu element induced EGFP protein circadian expression

After cell synchronization, EGFP fluorescence was recorded with real-time video microscopy. Individual cells included in the analysis are those that can be monitored during at least 48 consecutive hours and whose EGFP fluorescence intensity is 10% higher than the background. Values obtained in these selected cells were fitted by a cosinor equation. Example of data from two cells that could (right part) or could not (left part) be fitted by cosinor equation were given in *Figure 5D*. An example of a rhythmic fluorescence pattern over 48 consecutive hours in an individual cell was also given in *Video 1*. The percent of cells that could be fitted by cosinor equation with a R<sup>2</sup>>0.5 was then calculated for each cell line. This percent was significantly higher in IRAlu-egfp cell line compared to the Alu-egfp and egfp cell lines (*Figure 5D*) allowing to show that insertion of an IRAlu element in the 3'-UTR of egfp mRNA promoted EGFP protein circadian expression.

# Characterization of paraspeckle associated mRNA

### Analysis of Neat1 RNA pull-down by RNA sequencing

To preserve weak in situ interactions, GH4C1 cells were cross-linked. RNA pull-down was performed using two specific biotinylated oligonucleotides (S oligo 1 and S oligo 2) and a non-specific biotinylated oligonucleotide. Libraries were generated from the purified RNAs obtained with the two specific oligonucleotides but no library could be obtained with the non-specific oligonucleotide due to the too small quantity of material recovered. After sequencing of the two libraries, results were analyzed using the Tophat/Cufflinks pipeline (*Trapnell et al., 2012*). Only transcripts which exhibited values of fragment per kilobase per million of mapped reads (FPKM) higher than 1 were taken into account. We assessed the specificity of Neat 1 RNA pull-down by crossing the FPKM>1-limited lists obtained with the two specific oligonucleotides. 4268 genes (28% of expressed transcripts in GH4C1



**Figure 5.** IRAlu element induced egfp mRNA circadian nuclear retention and EGFP circadian cytoplasmic expression. (A) Rhythmic ratio of nuclear versus cytoplasmic egfp mRNA levels in Alu-egfp and IRAlu-egfp cell lines. Experimental values (n=3 for each cell line) can be adequately fitted (R<sup>2</sup>>0.55) with a non-linear cosinor equation in which the period value is set to 24 hr (see also *Figure 5—source data 1*). (B–C) Effects of paraspeckle disruption by Neat1 siRNA or Neat1 antisens oligonucleotides (ASO) on egfp mRNA nuclear retention in IRAlu-egfp cells. Nuclear and cytoplasmic egfp mRNA were quantified by qPCR in each condition and normalized to the relative amount of gapdh mRNA (n=2 for each condition). (B) Ratio of nuclear versus cytoplasmic egfp mRNA levels are compared between negative control and Neat1 siRNA or ASO \*p<0.05 \*\*p<0.01 (C) Loss of rhythmic egfp mRNA nuclear retention in IRAlu-egfp cells after Neat1 siRNA or ASO. Experimental values (n=2 for each condition) cannot be adequately fitted (R<sup>2</sup><0.55) with a non-linear cosinor equation in which the period value is set to 24 hr. (D) eGFP fluorescence was recorded with real-time video microscopy in individual cells and values were fitted by a cosinor equation. Shown are examples of data from two cells that could not or could be fitted by *Figure 5 continued on next page* 



# Effects of knockdown of Neat1 by Neat1 siRNA or Neat1 ASO.

(A) On Neat1 RNA levels: Neat1 RNA levels are reduced by about 40% as a result of siRNA or ASO treatment. \*\*\*\*p<0.0001 (**B**–**D**) On the Neat1 RNA circadian pattern: The pattern of Neat1 RNA expression is determined by RT-qPCR in cells treated with Neat1 siRNA, Neat1 ASO or negative control. Experimental values from negative control siRNA treatment expressed as a percent of the initial value obtained at ZT 0 can be adequately fitted (R<sup>2</sup>>0.55) with a non-linear cosinor equation in which the period value is set to 24 hr (**B**). By contrast, this is not the case for experimental values obtained after Neat1 siRNA (**C**) or Neat1 ASO (**D**) treatment.

#### Figure 5 continued

cosinor equation, in left and right panels respectively. The percent of cells that could be fitted by cosinor equation with a  $R^2 > 0.55$  was then calculated for each cell line. This percent was significantly higher in IRAlu-egfp cell line compared to the Alu-egfp and egfp cell lines. \*p<0.05 DOI: 10.7554/eLife.14837.013

The following source data and figure supplement are available for figure 5:

**Source data 1.** Cosinor analysis of the rhythmic ratio of nuclear versus cytoplasmic egfp mRNA levels in Alu-egfp and IRAlu-egfp cell lines.

DOI: 10.7554/eLife.14837.014

**Figure supplement 1.** Effects of knockdown of Neat1 by Neat1 siRNA or Neat1 ASO. DOI: 10.7554/eLife.14837.015

cells) were common to the two lists (*Figure 6—source data 1*). This represented 65% of the list obtained with S oligo 1 and 83% of the list obtained with S oligo 2.

#### Association of rhythmic transcripts with paraspeckles

Using publicly available datasets, we then tried to evaluate whether known cycling transcripts can be shown to be associated with paraspeckles. To allow comparison between gene lists generated from different species, namely rat and mouse, we used only official gene symbols (3928 genes in our 4268 gene list if we consider only official gene symbols). We first compared the mouse pituitary circadian transcriptome previously published (Hughes et al., 2007) with our data set of Neat1 RNA targets. It appeared that 68 genes out of 362 from the list of pituitary circadian transcripts (18.8%) were found to be included in our data set (Figure 6A). We then tried to assess whether genes known to be post-transcriptional cyclers are associated with paraspeckles. To this end, we compared the list of putative post-transcriptional cycling transcripts established by Menet et al. (Menet et al., 2012) in the mouse liver with our list of Neat1 RNA targets. Although these two lists have been established in two different tissues from two different species, it appeared that 182 genes out of 675 from the list of post-transcriptional circadian genes in the liver (27%) were found to be included in our data set (Figure 6A). In agreement with the assumption that rhythmic circadian genes are tissue specific, we found very few genes (24 genes) common to the list of Hughes et al. (<7%) (Hughes et al., 2007) and that of Menet et al. (<4%) (Menet et al., 2012). In spite of this reduced overlap, we found that 6 genes out of these 24 common genes (25%) were Neat1 RNA targets (Figure 6A).

#### Circadian nuclear retention of Neat1 RNA target genes

We then tested on a few genes whether Neat1 RNA targets as identified here can display a circadian nuclear retention. With this aim, we chose among Neat1 RNA targets, six genes that were included either in the list of Menet et al. (*Menet et al., 2012*) namely Calr, Evi5 and Maged1 or in the list of Hughes et al. (*Hughes et al., 2007*) namely Canx, Pcbp2 and Syne1 and one gene that was common to both lists namely Fkbp4 (*Figure 6A*). First we verified by qPCR the specific association of the seven selected genes with Neat1. To this end, we determined the enrichment in these seven mRNA obtained after Neat1 RNA pull-down with the two specific biotinylated complementary oligonucleotides that target Neat1 compared to the biotinylated irrelevant probe. As shown in *Figure 6—figure supplement 1*, the seven selected genes were significantly enriched after Neat1 RNA pull-down by

the two specific probes compared to the nonspecific probe. We further found that the seven selected genes displayed a rhythmic pattern of nuclear retention in IRAlu-egfp cells since values obtained for their relative mRNA ratio between nucleoplasm and cytoplasm over the time can be fitted by a sine wave curve with a R squared >0.55 (*Figure 6B, Figure 6—figure supplement 2, Figure 6—source data 2*). It may be noticed that the minimum value of the relative ratio nuclear / cytoplasmic mRNA occurred roughly 20 hr after the medium



Video 1. Rhythmic fluorescence pattern of EGFP expression over 48 consecutive hours in an individual cell from IRAlu-egfp cell line. DOI: 10.7554/eLife.14837.016



**Figure 6.** Characterization of paraspeckle-associated mRNA. (A) Venn diagram representation of the overlap between transcripts linked to the paraspeckle (Neat1 RNA targets) with transcripts known as rhythmic genes in the mouse pituitary gland (*Hughes et al., 2007*) and/or post-transcriptional rhythmic genes in mouse liver (*Menet et al., 2012*). To allow comparison between lists from two different species, lists of genes were restricted to official gene symbol lists. Shown are the seven genes that were selected for analysis of mRNA nuclear/cytoplasmic ratio; in bold the three *Figure 6 continued on next page* 



#### Figure 6 continued

genes illustrated in the figure. (B) Rhythmic ratio of nuclear versus cytoplasmic mRNA levels of three selected genes. Experimental values (n=2) can be adequately fitted ( $R^2$ >0.55) with a non-linear cosinor equation in which the period value is set to 24 hr (see also *Figure 6—source data 1*). (C–D) Loss of rhythmic ratio of nuclear versus cytoplasmic mRNA levels of 3 selected genes after Neat1 siRNA (C) or Neat1 ASO (D). Experimental values (n=2) cannot be adequately fitted ( $R^2$ <0.55) with a non-linear cosinor equation in which the period value is set to 24 hr. DOI: 10.7554/eLife.14837.017

The following source data and figure supplements are available for figure 6:

Source data 1. List of Neat1 RNA target genes. DOI: 10.7554/eLife.14837.018 Source data 2. Cosinor analysis of the relative ratio nuclear/cytoplasmic mRNA of seven Neat1 RNA targets in GH4C1 cells. DOI: 10.7554/eLife.14837.019 Figure supplement 1. Enrichment relative to input in seven selected mRNA after Neat1 RNA pull-down. DOI: 10.7554/eLife.14837.020 Figure supplement 2. Rhythmic ratio of nuclear versus cytoplasmic mRNA levels of four NEAT1 RNA target genes.

Figure supplement 2. Rhythmic ratio of nuclear versus cytoplasmic mRNA levels of four NEAT1 RNA target gene: DOI: 10.7554/eLife.14837.021

change, i.e. at a time when the levels of paraspeckle protein components and Neat1 RNA were the lowest (*Figure 3*).

# Loss of circadian nuclear retention of Neat1 RNA targets after paraspeckle disruption

To determine whether mRNA retention in paraspeckles can drive the oscillations of selected transcripts, we performed a loss of function experiment. In *Figure 6C–D* and *Figure 6—figure supplement 2*, it was shown that knockdown of Neat1 RNA, either by Neat1 siRNA (*Figure 6C*) or by Neat1 ASO (*Figure 6D*) disrupted the circadian pattern of the ratio of nuclear versus cytoplasmic levels of the seven selected mRNA as assessed by the inability to fit the values obtained with a sine wave curve with a R squared >0.55. These results showed that paraspeckles play a role in the rhythmic expression of theses selected genes.

# Discussion

# Features of paraspeckle bodies in GH4C1 cells

Paraspeckles are found in almost all of the cultured cell lines and primary cultures from tissues (Bond and Fox, 2009), except for embryonic stem cells (Chen and Carmichael, 2009). Four RNAbinding proteins, including three members of the Drosophila melanogaster behavior human splicing (DBHS) family proteins (NONO, PSPC1 and SFPQ) and RNA-binding motif protein 14 (RBM14) are classified as 'classical' paraspeckle protein components (Bond and Fox, 2009; Prasanth et al., 2005; Fox et al., 2002; Nakagawa and Hirose, 2012). Both endogenous and tagged forms of these proteins are found localized within the nucleoplasm as well as paraspeckles in mammalian cells. With a confocal fluorescent study, we report in a rat pituitary cell line that all four paraspeckle protein components are distributed both in the nucleoplasm and in subnuclear foci. Among these later, we identify paraspeckle nuclear bodies by overlapping side-by-side the staining of two paraspeckle protein components. This allows not only to provide evidence for the existence of paraspeckles in rat pituitary cell nuclei but also to show that as already reported in other cell types, paraspeckles are small, irregularly sized and unevenly distributed subnuclear bodies. Using RIP experiments, we further show that all four paraspeckle protein components are associated with the specific paraspeckle localized IncRNA Neat1. This later is further detected by RNA FISH. Confirming previous reports in other cell types, when identified by a combination of Neat1 RNA FISH and confocal microscopy, paraspeckles appear as round foci in rat pituitary cells. The diameter of such round foci is in the range of that estimated in previous studies (around 360 nm). However after use of a combination of Neat1 RNA FISH and STORM analysis, paraspeckles appear more likely as oblong structures with smaller dimensions reminiscent of the interchromatin granule-associated zone (IGAZ) that had been observed under electron microscope and had been reported to correspond to paraspeckles (Cardinale et al., 2007; Bond and Fox, 2009; Souquere et al., 2010). Width and height

# Figure 6—figure supplement 1



# Enrichment relative to input in seven selected mRNA after Neat1 RNA pulldown.

Enrichment relative to input obtained after using the two anti-sens specific (S oligo 1 and S oligo 2) or a non-specific (NS) oligonucleotide probes. \*\*\*p<0.001vs non-specific oligonucleotide probe.

# Figure 6—figure supplement 2



# Enrichment relative to input in seven selected mRNA after Neat1 RNA pulldown.

Enrichment relative to input obtained after using the two anti-sens specific (S oligo 1 and S oligo 2) or a non-specific (NS) oligonucleotide probes. \*\*\*p<0.001vs non-specific oligonucleotide probe.

measurements obtained after STORM analysis of Neat1 RNA FISH staining are smaller than the dimensions of IGAZ nuclear bodies established by an electron microscopy analysis in human and mouse cells (*Souquere et al., 2010*) suggesting that either IGAZ/paraspeckle nuclear bodies are smaller in rat than in mouse cells or alternatively and more likely, Neat1 RNA staining does not fill in all the nuclear bodies as already proposed (*Souquere et al., 2010*).

In a previous study, we reported that two paraspeckle proteins, namely NONO and SFPQ, display a circadian expression pattern in rat pituitary cells (Guillaumond et al., 2011) and we find here that it is also the case for the two other 'classical' paraspeckle proteins, PSPC1 and RBM14. Furthermore, all four proteins bind rhythmically to Neat1, suggesting that the pools of proteins involved in paraspeckle composition vary with time. In addition, the IncRNA Neat1 itself displays a circadian expression pattern and we further found using a pair of primers allowing the specific detection of the long form of Neat1 RNA, Neat1-2 (Figure 2-source data 1) that the expression of this form known to be sufficient for the formation of paraspeckle displays a circadian pattern (data not shown). The circadian expression pattern of Neat1 RNA is not the privilege of 'in vitro' pituitary cells since it is also found 'ex vivo' in several mouse tissues including the pituitary gland but also numerous other peripheral oscillators such as the spleen or the adrenal gland. This holds also true for the central clock, namely the suprachiasmatic nuclei. Finally, in GH4C1 cells, the circadian variations in Neat1 RNA levels are shown here to reflect rhythmic variations in paraspeckle number within the cells. It may then be assumed that the circadian expression pattern of Neat1 RNA described here in a number of mouse tissue also reflects variations in paraspeckle number although it cannot be excluded that paraspeckle size could also oscillate over the circadian period both in pituitary cells and in other oscillators.

# Association of IRAlu elements with paraspeckles

As previously reported by Chen and Carmichael (**Chen et al., 2008**), we found that a pair of IRAlus in the 3'-UTR of egfp can strongly silence EGFP expression in GH4C1 pituitary cell line and after fractionating cytoplasmic and nuclear RNAs from our two cell lines, we observed that IRAlus-containing egfp mRNA appear to be preferentially retained in the nucleus in comparison with those having a single Alu element. It then appeared that silencing EGFP expression mechanism relied on egfp mRNA nuclear retention.

To obtain evidence that nuclear retained IRAlus are associated with paraspeckle bodies, we immunoprecipitated RNA complexes containing paraspeckle protein PSPC1 and characterized the associated egfp mRNA. We found IRAlu-containing egfp mRNA to be associated with PSPC1 whereas Alu-containing egfp mRNA was not. While PSPC1 as well as NONO and SFPQ accumulate in paraspeckles, these proteins are also found elsewhere throughout the nucleoplasm and thus may have functions apart from paraspeckle retention. In contrast to paraspeckle-associated proteins that are multifunctional, NEAT1 may have only one function. We then developed a technique to characterize nuclear mRNA Neat1 targets using a Neat1 RNA pull-down approach that allowed us to show the association of IRAlu egfp mRNA with Neat1 providing a much more convincing evidence for a nuclear retention of IRAlu egfp mRNA by paraspeckles.

# IRAlu element-induced circadian expression of reporter gene

Given their presumed functions in gene expression through corresponding mRNA nuclear retention and since they display a circadian expression pattern, it may be assumed that paraspeckle bodies can rhythmically retain RNAs in the nucleus leading to a rhythmic expression of the corresponding gene. In keeping with this view, we showed that IRAlu elements inserted in 3'-UTR of egfp reporter mRNA allow for its circadian retention within the nucleus and consequently for EGFP protein circadian expression. However, it may be noticed that egfp mRNA displayed also a circadian nuclear retention when a single antisens Alu element was inserted in its 3'-UTR but in this case the variations of the ratio over the circadian period was highly reduced in contrast to the high amplitude circadian variations of the nuclear versus cytoplasmic ratio observed in IRAlu-egfp cell line. These small amplitude variations of the ratio between nuclear and cytoplasmic mRNA contents in Alu-egfp cell line were further accompanied by a weak number of cells expressing a circadian pattern of EGFP expression. Since antisens Alu elements within the 3'-UTR of mRNAs can be the target of several RNAbinding proteins (*Kelley et al., 2014; Zarnack et al., 2013*) it may be speculated that some of these RNA-binding proteins display a rhythmic expression pattern accounting for the weak EGFP circadian expression observed in Alu-egfp cell line. Among the identified RNA-binding proteins that target antisens Alu elements within the 3'-UTR of mRNAs, HNRNPC was shown to display a circadian expression pattern in several tissues including the mouse pituitary (see CIRCA: http://circadb.hoge-neschlab.org/). This is not the case for HNRNPU and STAUFEN 1 (STAU1)-mediated messenger RNA decay we did not find to display a circadian expression pattern in GH4C1 cells (*Guillaumond et al., 2011*) and (data not shown) and are then unlikely involved in circadian EGFP expression in Alu-egfp cell line. The mechanism involved in the circadian nuclear retention of Alu-containing egfp mRNA remains anyway to be determined.

# Contribution of paraspeckle nuclear retention to circadian gene expression

Nuclear retention of egfp mRNA in IRAlu-egfp cells is no more cycling when paraspeckles were disrupted assuming that the nuclear retention of the reporter gene by paraspeckles contributes to cycling. To determine whether this paraspeckle impact on rhythmic gene expression occurs 'in vivo' on a much larger scale, we characterized in our pituitary cells the list of the putative Neat1 RNA target genes and we analyzed in publicly available datasets the known cycling transcripts that are included in this Neat1 RNA target list. Whereas it is well known that rhythmic circadian genes are tissue specific and that it is unfavorable to compare datasets obtained in two different species we found a robust overlap of our gene list with circadian genes in the mouse pituitary (Hughes et al., 2007) and a closer overlap with post-transcriptional circadian genes in the mouse liver (Menet et al., 2012). Furthermore, on a few circadian genes common to either of these lists or both, we showed a loss of rhythmic pattern when paraspeckles were disrupted assuming that paraspeckle retention plays a relevant role in circadian gene expression. It is now widely admitted that post-transcriptional circadian regulation plays a major role in determining oscillations at the mRNA level. In agreement with this view, the new mechanism described here suggests that circadian mRNA oscillations could be post-transcriptionally controlled through rhythmic nuclear retention by paraspeckle nuclear bodies.

# **Materials and methods**

# Cell line culture and preparation of stably transfected cell lines

GH4C1 cells, a rat pituitary somatolactotroph line, were obtained in 2012 from ATCC (CCL-82.2, lot number: 58945448, Molsheim, France) with certificate analysis and were confirmed to be free of mycoplasma (MycoAlert, Lonza, Levallois-Perret, France). They were grown in HamF10 medium supplemented with 15% horse serum and 2% fetal calf serum. To synchronize cells between themselves, GH4C1 cells were transferred to fresh medium and were harvested after this fresh medium replacement (T0) every 4 hr from T2 to T30-38. For the generation of stable GH4C1 cell lines, cells were transfected with the plasmid constructs expressing a neomycin resistance gene by Lipofectamine Plus (Invitrogen, Cergy Pontoise, F). Cells were selected with 250 µg/ml G418 (Invitrogen) beginning 48 hr after transfection.

# **Plasmid constructs**

Alu and IRAlus elements from Nicn1 were cloned by PCR from genomic DNA obtained from Hela cells. Sequences were then inserted in the pEGFP-C1 (Clontech, Mountain View, CA) at the BgIII and HindIII sites as described by Chen et al. (*Chen and Carmichael, 2008*).

# Knockdown of Neat1

Predesigned siRNAs targeting rat Neat1 were purchased from GeneCust (siRNA Set, GeneCust, Dudelange, Luxembourg). IRAlu-egfp cells plated in 6-well dishes were transfected with 500 pmol of a pool of four siRNA (125 pmol each, see *Figure 2—source data 1* for the sequences) using Lipo-fectamine 2000 (Life Technologies, Saint Aubin, France). As controls, 500 pmol of non-targeting siRNA from GeneCust (negative control siRNA, see *Figure 2—source data 1* for the sequences) were used. Cells were harvested 48 hr after transfection.

The antisense oligonucleotides (ASO) used for Neat1 knockdown experiments were phosphorothioate modified at their backbone to increase their stability. Two Neat1 ASO were pooled that were targeted to regions common to NEAT1\_1 and NEAT1\_2 isoforms (see *Figure 2—source data 1* for the sequences). IRAlu-egfp cells plated in 6-well dishes were transfected with 300 pmol of the pool of the two ASO (150 pmol each, see *Figure 2—source data 1* for the sequences) using Lipofectamine 2000 (Life Technologies). Cells were harvested 48 hr after transfection.

# **RNA** expression analysis

Nuclear, cytoplasmic or total RNA was prepared from GH4C1 cells using an RNA XS kit (Macherey Nagel, Hoerdt, France). Total RNA purification was performed on 24-well cell dishes. Nuclear and cytoplasmic RNA isolation was performed using 10 cm cell dishes that were rinsed twice with ice-cold PBS, incubated in 1 ml of ice-cold cell lysis buffer A (10 mmol/L Tris pH 7.4, 3 mmol/L MgCl2, 10 mmol/L NaCl and 0.5% NP-40). Nuclei and cytoplasma were separated by centrifugation (500 x *g* for 5 min). One-sixth of the supernatant was used to prepare cytoplasmic RNA. To obtain pure nuclear RNA, the nuclear pellets were subjected to two additional washes with 1 ml lysis buffer A and were then extracted with XS kit reagent. Total RNA (500 ng) was then used for cDNA synthesis performed with a High Capacity RNA to cDNA kit (Applied Biosystem, Courtaboeuf, France). Real-time PCR was performed on a 7500 fast Real-Time qPCR system (Applied Biosystems) using Fast SYBR Green mix (Applied Biosystems). The sequences of the primers used in qPCR are given in *Figure 2—source data 1*. mRNA accumulation was normalized relative to Gapdh mRNA levels.

# Western-blot analysis

Total or nuclear protein extracts prepared as previously described (**Becquet et al., 2001**) from confluent GH4C1 cells grown in 10 cm dishes, were submitted to Western-blot analysis as previously described (**Guillaumond et al., 2011**) with polyclonal primary antibodies raised against PSPC1 (1:1000, sc-84576, Santa Cruz Biotechnology, Heidelberg, Germany), RBM14 (1  $\mu$ g/ml, Abcam, Paris, France), ATF2 (1/1000, sc-187, Santa Cruz Biotechnology), GFP (1/1000, A6455, Molecular Probes, Paisley, UK) and monoclonal primary antibody raised against  $\alpha$ -Tubulin (1/1000, T 6199, Sigma-Aldrich, Saint-Louis, USA). The ratio of PSPC1 and RBM14 to ATF2, a constitutive transcription factor, and the ratio of GFP to tubulin were determined by densitometry using ImageJ software (National Institutes of Health, USA).

# **RNA** immunoprecipitation (RIP) experiments

GH4C1 cells grown in 10 cm dishes were rinsed two times with 5 ml cold phosphate buffer saline (PBS). Cells were then harvested by scraping in ice-cold PBS and transferred to a centrifuge tube. After centrifugation (2500 x g for 5 min) cells were pelleted and suspended in 100  $\mu$ l of Polysome lysis buffer (PLB; 10 mM HEPES, pH 7.0, 0.1M KCl, 5 mM MgCl2, 25 mM EDTA, 0.5% NP40, 1 mM DTT, 100 U/ml RNAse OUT and complete protease inhibitor cocktail). After mixing by pipetting up and down, cells were kept on ice for 5 min to allow the hypotonic PLB buffer to swell the cells. The cell lysate was then aliquoted and stored at  $-80^{\circ}$ C.

Cell lysate was centrifuged at 14,000 x g for 10 min at 4°C and diluted 1/100 in NET2 buffer (NET2 buffer corresponded to NT2 buffer: 50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 1 mM MgCl2 and 0.05% NP40 added with 1 mM DTT, 20 mM EDTA, 200 U/ml RNAse Out). An aliquot of diluted cell lysate was removed and represented the starting material or 'input' which was processed along-side the immunoprecipitation to compare with immunoprecipitated mRNAs at the end. RIP experiments were performed overnight at 4°C on diluted cell lysate with antibodies to NONO (ab45359, Abcam), SFPQ (ab38148, Abcam), PSPC1 (SAB4200068, Sigma-Aldrich, Saint-Quentin Fallavier, France) and RBM14 (ab70636, Abcam) or non-specific rabbit polyclonal antibody (anti-Furin, sc-20801, Santa-Cruz Biotechnology). After incubation was completed, 15 µl of Magna ChIP protein A magnetic beads (16–661, Millipore, Molsheim, France) were added for 1h at 4°C. Beads were washed 6 times with cold NT2 buffer and treated by proteinase K for 30 min at 55°C. RNA eluted from beads was purified using Nucleospin RNA XS (Macherey-Nagel) and processed for cDNA synthesis using a High Capacity RNA to cDNA kit (Applied Biosystems).

#### Neat1 RNA pull-down

Neat1 RNA pull-down is a hybridization-based strategy that uses complementary oligonucleotides to purify Neat1 RNA together with its targets from reversibly cross-linked extracts. In cross-linked extracts, it is expected that some regions of the RNA will be more accessible for hybridization than others due in particular to secondary structure. To design oligonucleotides that target these regions and then can hybridize specifically to Neat1 RNA, we modeled the secondary structure of Neat1 RNA using bioinformatics (MaxExpect software [*Lu et al., 2009*]). Two antisense DNA oligonucleotide probes that target accessible regions of the lncRNA Neat1were designed and used for Neat1 RNA specific pull-down (*Figure 4—figure supplement 1*). One biotynylated irrelevant probe was used for Neat1 RNA non-specific pull-down. All three probes were biotinylated at the 3' end (*Figure 2—source data 1*).

Briefly, 10 cm cell dishes were incubated in 1 ml of ice-cold cell lysis buffer A as described above. Nuclei were scraped and separated by centrifugation (500 x g for 5 min). The nuclear pellets were then fixed with 1% paraformaldehyde in PBS for 10 min at room temperature. Crosslinking was then quenched with 1.25 M glycine for 5 min. Cross-linked nuclei were rinsed two times again with PBS and pelleted at 500 x g. Nuclear pellets were stored at -80°C. To prepare lysates, nuclear pellets were suspended in lysis buffer (50 mM Tris, pH 7.0, 10 mM EDTA, 1% SDS added with a protease inhibitor cocktail and RNAse-Out), and sonicated using BioruptorPlus (Diagenode, Seraing, Belgium) by 2 pulses of 30 s allowing complete lysate solubilization. RNA was in the size range of 400 to 2000 nucleotides. Nuclear lysates were diluted V/V in hybridization buffer (750 mM NaCl, 1% SDS, 50 mM Tris, pH 7.0, 1 mM EDTA, 15% formamide). The two specific or the non-specific probes (100 pmol) were added to 1.2 ml of diluted lysate, which was mixed by end-to-end rotation at room temperature 4 to 6 hr. Streptavidin-magnetic C1 beads (Dynabeads My OneStreptavidin C1- Invitrogen Life Technologies) were added to hybridization reaction (50  $\mu$ l per 100 pmol of probes) and the whole reaction was mixed overnight at room temperature. Beads-biotin-probes-RNA adducts were captured by magnets (Invitrogen) and washed five times with a wash buffer (2×SSC, 0.5% SDS). After the last wash, buffer was removed carefully. For RNA elution, beads were suspended in 100  $\mu$ l RNA proteinase K buffer (100 mM NaCl, 10 mM Tris, pH 7.0, 1 mM EDTA, 0.5% SDS) and 100 μg proteinase K (Ambion). After incubation at 45°C for 45 min, RNA was isolated using NucleoSpinRNA XS (Macherey-Nagel). Eluted RNA was subject to RNA-sequencing or RT-qPCR for the detection of enriched transcripts.

#### **RNA** sequencing

The construction of Illumina DNA libraries from RNA pools obtained with each of the two specific oligonucleotides and the sequencing were performed by Helixio (Clermont-Ferrand, France). Libraries were prepared with Illumina TruSeq Stranded mRNA Sample Preparation kit. Sequencing was done on Illumina NextSeq 500 (60 million reads per sample on average).

Analyses were performed on a local instance of Galaxy (Afgan et al., 2016). Paired-end reads (75bp) were aligned to the Rat reference genome (Rnor\_6.0.80, Ensembl) using TopHat2 v2.209 with default parameters (Kim et al., 2013). We further used Cuffflinks v2.2.1 (Trapnell et al., 2012) to quantify mRNA level and fragment per kilobase per million of mapped reads (FPKM) were calculated. Consistent with prior studies (Nagaraj et al., 2011; Hart et al., 2013), the log2 (FPKM) distribution shows a primary peak of high expression genes, with a long left shoulder of low-expression transcripts. Therefore, only transcripts with FPKM values higher than 1 were taken into account. In order to assess the specificity of the NEAT1 RNA pull-down we crossed the results obtain with the two oligonucleotides (see Neat1 RNA pull-down) to generate the list of transcripts associated to the paraspeckles.

All the RNA sequencing data are available at Gene Expression Omnibus (GEO) (accession no. GSE81972).

#### Multiple-label fluorescence immunocytochemistry

GH4C1 cells were grown on glass coverslips coated with poly-ornithine. After incubation in 30% normal blocking serum containing 0.2% Triton X-100 (30 min), coverslips were transferred to primary antibody or a mixture of primary antibodies (overnight, 4°C) and, after rinsing, to fluorescent secondary antibody or a mixture of fluorescent secondary antibodies (2 hr, 22°C). Nuclei were stained with Hoechst dye (1  $\mu$ mol/ml, 5 min). All dilutions and rinsing between each step of the procedures were made in Tris buffer added with 1% normal serum.

NONO, SFPQ, PSPC1, RBM14 were detected using the respective polyclonal rabbit antibodies: ab45359 (1: 1000) Abcam, ab38148 (1: 1000) Abcam, HPA038904 (1: 1000) Sigma-Aldrich, ab70636 (1: 1000) Abcam.

After mounting with PBS containing 50% glycerol, the reacted sections were kept in darkness at  $4^{\circ}$ C until examination.

### **RNA-FISH**

To detect Neat1 RNA, GH4C1 cells grown on glass coverslips coated with poly-ornithine were fixed in 3.6% formaldehyde. Hybridization was carried out using Custom Stellaris FISH Probes (Biosearch Technologies, Novato, USA). Probes used are Neat1 probes labeled with Quasar 570 Dye or Neat1 probes labeled with Quasar 670 for STORM analysis. Nuclei of the cells were counterstained by Hoechst solution (5 ng/ml).

# **Confocal microscopy analysis**

The confocal image acquisition was performed on a Zeiss LSM780 confocal microscope equipped with a 63X 1.4 N.A. oil immersion objective. Dyes used in either immunocytochemistry (ICC) or RNA-Fish were imaged using the appropriate wavelength for optimal dye excitation, i.e. 405 nm for Hoechst, 561 nm for Cy3 and Quasar 570, and 633 nm for both Cy5 and Quasar 670. The spectral detection of emitted fluorescence were set as follows: 420–480 nm for Hoechst, 560–670 nm for Cy3 or Quasar 570 and 640–750 nm for Cy5 or Quasar 670. Three dimensional z-stacks were collected automatically as frame by frame sequential image series. To enhance resolution, the ICC images were deconvolved based on a theoretical point spread function (PSF) using AutoQuantX (Media Cybernetics, Rockville, USA).

# **STORM** microscopy

Prior to observation, samples were mounted in imaging buffer (50 mM Tris pH 8.0, 20 mM NaCl, 0.5 mg/ml glucose oxydase, 40 µg/ml catalase, 10% w/v glucose, 10 mM mercaptoethylamine). Imaging was performed with a Nikon Storm System (N-Storm, Nikon France S.A, Champigny sur Marne, France) equipped with a 100X 1.49 N.A. oil immersion objective. A reference conventional fluorescence image of the cell of interest was acquired before recording storm sequence of images. For super resolution image acquisition, the sample stained with Quasar 670 dye was excited by a 647 nm imaging laser. Illumination with 405 nm light was gradually increased along recording to reactivate Quasar 670 blinking. The fluorescence emission was separated from the excitation light by adapted dichroic mirror (660 nm) and filters (692/4 nm) and collected on an IXon EMCCD camera (Andor Technology, Belfast, UK). Typically, dSTORM images were reconstructed from 40000 frames using NIS-software (Nikon).

# Real-time monitoring of Egfp fluorescence by video-microscopy and mathematical analysis

For real-time monitoring of intracellular fluorescence circadian oscillations,  $4 \times 10^4$  Alu-egfp or IRAlu-egfp GH4C1 cells were seeded on glass-bottom 24 well plates (coated with poly- ornithine). The next day, the medium was changed with a growth medium containing 100 nM forskolin to synchronize the cellular circadian oscillators. 20 min later, the medium was changed again with a DMEM without phenol red medium containing 100 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) and 100 mM pyruvate. The next day, plates were transferred to an inverted microscope (Axiovert-200, Zeiss) equipped with an environmental chamber, the temperature of which was set to  $37^{\circ}$ C and CO<sub>2</sub> regulated to 5%. Egfp was excited through a 475/40 band-pass filter. An EMCCD camera (Rolera EM-C2, Q-Imaging) was used to collect emitted fluorescence at 530/50 nm. Time-lapse recordings of multiple regions in each well were realized over 96 hr with 1 picture every 30 min (Metamorph, Roper Scientific), using a 40X objective. The data were processed for tracking and measuring fluorescence fluctuations of individual cells with the CGE (Circadian Gene Expression) Plugging of ImageJ (**Sage et al., 2010**). Cells retained for analysis fulfilled two criteria: they were

detected for at least 48 hr and their mean fluorescent level was 10% higher than mean level of the background.

# Paraspeckle number quantification

After processing of the cells for FISH of Neat1 RNA, quantification of paraspeckle bodies was performed on confocal images using a 40X objective. At each time point analyzed, 20 to 35 fields acquired in four wells from two different experiments were counted to estimate the total paraspeckle number per well of 100000 cells and to estimate among cells which expressed paraspeckles the mean number per cell.

# Cosinor and statistical analysis

Data from real time monitored cells or mean experimental values ( $\pm$  SEM) expressed as a percent of initial value were fitted (Prism4 software, GraphPad Software, Inc.) by a non-linear sine wave equation to find the set of parameters that gives the least-squared distance between the data and the equation:  $y(t) = B + A^{*}exp(-t)^{*}sin(2^{*}pi^{*}t/T + P)$  where *B* is the Baseline, *A* is the Amplitude, *T* is the period and *P* is the Phase-shift. Goodness-of-fit was quantified using R squared, experimental values being considered well fitted by cosinor regression when the R squared was higher than 0.55. A statistically significant circadian oscillation was considered if the 95% confidence interval for the amplitude did not include the zero value (zero-amplitude test) (Yang et al., 2009) (Kavcic et al., 2011).

Significant differences between groups were determined using either ordinary one-way ANOVA or unpaired t test (Prism4 software). Values were considered significantly different for p<0.05 (\*).

# Acknowledgements

The authors gratefully acknowledge Nikon France SA (Champigny sur Marne, France) to provide a freely available Nikon Storm System.

# **Additional information**

#### Funding

Funder	Grant reference number	Author
Pfizer	SFE award - Research price 2014	Anne-Marie François-Bellan

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

#### Author contributions

MT, DB, M-PB, Performed experiments, Analyzed data and interpreted results of experiments, Prepared figures, Edited and revised manuscript; SG, Performed cell culture, Established stably transfected cell lines, Performed plasmid constructs, Analysis and interpretation of data; BB, Performed cell cultures, Transfection experiments, western blot analysis and qPCR analysis; MM, Established stably transfected cell lines and performed FISH experiments, Analysis and interpretation of data; J-LF, Conception and design of research, Performed experiments, Analyzed data and interpreted results of experiments, Prepared figures, Edited and revised manuscript; A-MF-B, Conception and design of research, Performed experiments, Analyzed data of interpreted results of experiments, Prepared figures, Drafted manuscript, Edited and revised manuscript

#### Author ORCIDs

Jean-Louis Franc, Dhttp://orcid.org/0000-0002-2900-5468 Anne-Marie François-Bellan, Dhttp://orcid.org/0000-0002-3278-4642

Databasa liconso

# **Additional files**

#### Major datasets

#### The following dataset was generated:

Author(s)	Year	Dataset title	Dataset URL	and accessibility information
Torres M, Becquet D, Blanchard M, Guillen S, Boyer B, Moreno M, Franc J, François-Bellan A	2016	Determination of the RNA linked to the paraspeckles in the GH4C1 cell line	http://www.ncbi.nlm.nih. gov/geo/query/acc.cgi? acc=GSE81972	Publicly available at the NCBI Gene Expression Omnibus (accession no: GSE81972)

#### The following previously published dataset was used:

Author(s)	Year	Dataset title	Dataset URL	Database, license, and accessibility information
Menet JS, Rodri- guez J, Abruzzi KC, Rosbash M	2012	Nascent-Seq Reveals Novel Features of Mouse Circadian Transcriptional Regulation	http://www.ncbi.nlm.nih. gov/geo/query/acc.cgi? acc=GSE36916	Publicly available at the NCBI Gene Expression Omnibus (accession no: GSE36916)

### References

- Afgan E, Baker D, van den Beek M, Blankenberg D, Bouvier D, Čech M, Chilton J, Clements D, Coraor N, Eberhard C, Grüning B, Guerler A, Hillman-Jackson J, Von Kuster G, Rasche E, Soranzo N, Turaga N, Taylor J, Nekrutenko A, Goecks J. 2016. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Research* 44:W3–W10. doi: 10.1093/nar/gkw343
- Asher G, Schibler U. 2011. Crosstalk between components of circadian and metabolic cycles in mammals. *Cell* Metabolism **13**:125–137. doi: 10.1016/j.cmet.2011.01.006
- **Becquet D**, Guillaumond F, Bosler O, François-Bellan AM. 2001. Long-term variations of AP-1 composition after CRH stimulation: consequence on POMC gene regulation. *Molecular and Cellular Endocrinology* **175**:93–100. doi: 10.1016/S0303-7207(01)00393-8
- Becquet D, Boyer B, Rasolonjanahary R, Brue T, Guillen S, Moreno M, Franc JL, François-Bellan AM. 2014. Evidence for an internal and functional circadian clock in rat pituitary cells. *Molecular and Cellular Endocrinology* 382:888–898. doi: 10.1016/j.mce.2013.11.004
- Bond CS, Fox AH. 2009. Paraspeckles: nuclear bodies built on long noncoding RNA. The Journal of Cell Biology 186:637–644. doi: 10.1083/jcb.200906113
- Cardinale S, Cisterna B, Bonetti P, Aringhieri C, Biggiogera M, Barabino SM. 2007. Subnuclear localization and dynamics of the Pre-mRNA 3' end processing factor mammalian cleavage factor I 68-kDa subunit. *Molecular Biology of the Cell* 18:1282–1292. doi: 10.1091/mbc.E06-09-0846
- Chen LL, Carmichael GG. 2008. Gene regulation by SINES and inosines: biological consequences of A-to-I editing of Alu element inverted repeats. *Cell Cycle* **7**:3294–3301. doi: 10.4161/cc.7.21.6927
- Chen LL, DeCerbo JN, Carmichael GG. 2008. Alu element-mediated gene silencing. *The EMBO Journal* **27**: 1694–1705. doi: 10.1038/emboj.2008.94
- Chen LL, Carmichael GG. 2009. Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA. *Molecular Cell* **35**:467–478. doi: 10.1016/j. molcel.2009.06.027
- Clemson CM, Hutchinson JN, Sara SA, Ensminger AW, Fox AH, Chess A, Lawrence JB. 2009. An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. *Molecular Cell* 33: 717–726. doi: 10.1016/j.molcel.2009.01.026
- Fox AH, Lam YW, Leung AK, Lyon CE, Andersen J, Mann M, Lamond AI. 2002. Paraspeckles: a novel nuclear domain. Current Biology 12:13–25. doi: 10.1016/S0960-9822(01)00632-7
- Guillaumond F, Boyer B, Becquet D, Guillen S, Kuhn L, Garin J, Belghazi M, Bosler O, Franc JL, François-Bellan AM. 2011. Chromatin remodeling as a mechanism for circadian prolactin transcription: rhythmic NONO and SFPQ recruitment to HLTF. *The FASEB Journal* **25**:2740–2756. doi: 10.1096/fj.10-178616
- **Guru SC**, Agarwal SK, Manickam P, Olufemi SE, Crabtree JS, Weisemann JM, Kester MB, Kim YS, Wang Y, Emmert-Buck MR, Liotta LA, Spiegel AM, Boguski MS, Roe BA, Collins FS, Marx SJ, Burns L, Chandrasekharappa SC. 1997. A transcript map for the 2.8-Mb region containing the multiple endocrine neoplasia type 1 locus. *Genome Research* **7**:725–735. doi: 10.1101/gr.7.7.725
- Hart T, Komori HK, LaMere S, Podshivalova K, Salomon DR. 2013. Finding the active genes in deep RNA-seq gene expression studies. *BMC Genomics* **14**:778. doi: 10.1186/1471-2164-14-778

- Hughes M, Deharo L, Pulivarthy SR, Gu J, Hayes K, Panda S, Hogenesch JB. 2007. High-resolution time course analysis of gene expression from pituitary. *Cold Spring Harbor Symposia on Quantitative Biology* **72**:381–386. doi: 10.1101/sqb.2007.72.011
- Hurley JM, Dasgupta A, Emerson JM, Zhou X, Ringelberg CS, Knabe N, Lipzen AM, Lindquist EA, Daum CG, Barry KW, Grigoriev IV SKM, Galagan JE, Bell-Pedersen D, Freitag M, Cheng C, Loros JJ, Dunlap JC. 2014. Analysis of clock-regulated genes in Neurospora reveals widespread posttranscriptional control of metabolic potential. PNAS 11:16995–17002. doi: 10.1073/pnas.1418963111
- Hutchinson JN, Ensminger AW, Clemson CM, Lynch CR, Lawrence JB, Chess A. 2007. A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. *BMC Genomics* **8**:39. doi: 10.1186/1471-2164-8-39
- Kavcic P, Rojc B, Dolenc-Groselj L, Claustrat B, Fujs K, Poljak M, Kavčič P. 2011. The impact of sleep deprivation and nighttime light exposure on clock gene expression in humans. *Croatian Medical Journal* 52:594–603. doi: 10.3325/cmj.2011.52.594
- Kelley DR, Hendrickson DG, Tenen D, Rinn JL. 2014. Transposable elements modulate human RNA abundance and splicing via specific RNA-protein interactions. *Genome Biology* **15**:537. doi: 10.1186/s13059-014-0537-5
- Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology* 14:R36. doi: 10. 1186/gb-2013-14-4-r36
- Koike N, Yoo SH, Huang HC, Kumar V, Lee C, Kim TK, Takahashi JS. 2012. Transcriptional architecture and chromatin landscape of the core circadian clock in mammals. *Science* **338**:349–354. doi: 10.1126/science. 1226339
- Kretz M, Siprashvili Z, Chu C, Webster DE, Zehnder A, Qu K, Lee CS, Flockhart RJ, Groff AF, Chow J, Johnston D, Kim GE, Spitale RC, Flynn RA, Zheng GX, Aiyer S, Raj A, Rinn JL, Chang HY, Khavari PA. 2013. Control of somatic tissue differentiation by the long non-coding RNA TINCR. *Nature* **493**:231–235. doi: 10.1038/ nature11661
- Lim C, Allada R. 2013. Emerging roles for post-transcriptional regulation in circadian clocks. *Nature Neuroscience* **16**:1544–1550. doi: 10.1038/nn.3543
- Lu ZJ, Gloor JW, Mathews DH. 2009. Improved RNA secondary structure prediction by maximizing expected pair accuracy. RNA 15:1805–1813. doi: 10.1261/rna.1643609
- Menet JS, Rodriguez J, Abruzzi KC, Rosbash M. 2012. Nascent-Seq reveals novel features of mouse circadian transcriptional regulation. *eLife* 1:e00011. doi: 10.7554/eLife.00011
- Nagaraj N, Wisniewski JR, Geiger T, Cox J, Kircher M, Kelso J, Pääbo S, Mann M. 2011. Deep proteome and transcriptome mapping of a human cancer cell line. *Molecular Systems Biology* 7:548. doi: 10.1038/msb.2011. 81
- Nakagawa S, Hirose T. 2012. Paraspeckle nuclear bodies–useful uselessness? *Cellular and Molecular Life Sciences* 69:3027–3036. doi: 10.1007/s00018-012-0973-x
- Partch CL, Green CB, Takahashi JS. 2014. Molecular architecture of the mammalian circadian clock. Trends in Cell Biology 24:90–99. doi: 10.1016/j.tcb.2013.07.002
- **Prasanth KV**, Prasanth SG, Xuan Z, Hearn S, Freier SM, Bennett CF, Zhang MQ, Spector DL. 2005. Regulating gene expression through RNA nuclear retention. *Cell* **123**:249–263. doi: 10.1016/j.cell.2005.08.033
- Rey G, Cesbron F, Rougemont J, Reinke H, Brunner M, Naef F. 2011. Genome-wide and phase-specific DNAbinding rhythms of BMAL1 control circadian output functions in mouse liver. *PLoS Biology* 9:e1000595. doi: 10. 1371/journal.pbio.1000595
- Sage D, Unser M, Salmon P, Dibner C. 2010. A software solution for recording circadian oscillator features in time-lapse live cell microscopy. *Cell Division* 5:17. doi: 10.1186/1747-1028-5-17

Sasaki YT, Hirose T. 2009. How to build a paraspeckle. Genome Biology 10:227. doi: 10.1186/gb-2009-10-7-227

- Simon MD. 2013. Capture hybridization analysis of RNA targets (CHART). Current Protocols in Molecular Biology Chapter 21:25. doi: 10.1002/0471142727.mb2125s101
- **Souquere S**, Beauclair G, Harper F, Fox A, Pierron G. 2010. Highly ordered spatial organization of the structural long noncoding NEAT1 RNAs within paraspeckle nuclear bodies. *Molecular Biology of the Cell* **21**:4020–4027. doi: 10.1091/mbc.E10-08-0690
- Sunwoo H, Dinger ME, Wilusz JE, Amaral PP, Mattick JS, Spector DL. 2009. MEN epsilon/beta nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. *Genome Research* 19:347–359. doi: 10.1101/gr.087775.108
- Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols* **7**:562–578. doi: 10.1038/nprot.2012.016
- Wang D, Liang X, Chen X, Guo J. 2013. Ribonucleoprotein complexes that control circadian clocks. International Journal of Molecular Sciences 14:9018–9036. doi: 10.3390/ijms14059018
- West JA, Davis CP, Sunwoo H, Simon MD, Sadreyev RI, Wang PI, Tolstorukov MY, Kingston RE. 2014. The long noncoding RNAs NEAT1 and MALAT1 bind active chromatin sites. *Molecular Cell* 55:791–802. doi: 10.1016/j. molcel.2014.07.012
- Yang X, Zhang YK, Esterly N, Klaassen CD, Wan YJ. 2009. Gender disparity of hepatic lipid homoeostasis regulated by the circadian clock. *Journal of Biochemistry* **145**:609–623. doi: 10.1093/jb/mvp018
- Zarnack K, König J, Tajnik M, Martincorena I, Eustermann S, Stévant I, Reyes A, Anders S, Luscombe NM, Ule J. 2013. Direct competition between hnRNP C and U2AF65 protects the transcriptome from the exonization of Alu elements. *Cell* **152**:453–466. doi: 10.1016/j.cell.2012.12.023



Zhang EE, Kay SA. 2010. Clocks not winding down: unravelling circadian networks. Nature Reviews Molecular Cell Biology **11**:764–776. doi: 10.1038/nrm2995 EXTRA VIEW

Taylor & Francis

ARTICLE HISTORY Received 21 November 2016

rhythm; nuclear RNA

**KEYWORDS** 

Revised 21 December 2016 Accepted 22 December 2016

circadian oscillators: circadian

retention; paraspeckle; post-

transcriptional mechanism

Check for updates

# Paraspeckles as rhythmic nuclear mRNA anchorages responsible for circadian gene expression

Manon Torres<sup>†</sup>, Denis Becquet<sup>†</sup>, Marie-Pierre Blanchard, Séverine Guillen, Bénédicte Boyer, Mathias Moreno, Jean-Louis Franc, and Anne-Marie François-Bellan D

Aix Marseille Université, CNRS, CRN2M-UMR7286, Faculté de Médecine Nord, Marseille, France

#### ABSTRACT

Circadian clocks regulate rhythmic gene expression levels by means of mRNA oscillations that are mainly driven by post-transcriptional regulation. We identified a new post-transcriptional mechanism, which involves nuclear bodies called paraspeckles. Major components of paraspeckles including the long noncoding RNA Neat1, which is the structural component, and its major protein partners, as well as the number of paraspeckles, follow a circadian pattern in pituitary cells. Paraspeckles are known to retain within the nucleus RNAs containing inverted repeats of Alu sequences. We showed that a reporter gene in which these RNA duplex elements were inserted in the 3'-UTR region displayed a circadian expression. Moreover, circadian endogenous mRNA associated with paraspeckles lost their circadian pattern when paraspeckles were disrupted. This work not only highlights a new paraspeckle-based post-transcriptional mechanism involved in circadian gene expression but also provides the list of all mRNA associated with paraspeckles in the nucleus of pituitary cells.

Most organisms have built-on time-measuring devices that are commonly known as circadian clocks. These circadian clocks allow them to anticipate the time of day and hence to temporally organize behavior as well as physiologic and biochemical processes. They therefore play a key role as adaptive mechanisms to permanent changes in the environment necessary for the individual survival and the sustainability of the species. In living systems ranging from bacteria to humans,<sup>1,2</sup> circadian rhythms are generated endogenously through genetic control<sup>3</sup> and regulate vital aspects of the organism physiology, from sleeping and waking to neurotransmitter secretion and cellular metabolism. At the center of these rhythms is the circadian clock machinery that consists in a transcription-translation feedback system regulated by a group of genes that oscillate in a circadian manner, the socalled *clock genes*. In mammals, the circadian system is hierarchically organized. Indeed, while molecular oscillations occur in most cells and tissues of the body,

a central structure, the suprachiasmatic nucleus (SCN) of the hypothalamus, functions as the master regulator to synchronize the phase of the other slave oscillating tissues.<sup>4,5</sup>

Briefly, the molecular mechanism that generates circadian rhythms involves the interacting positive and negative feedback loops of transcriptional or translational processes of clock genes (Fig. 1).<sup>1,6</sup> In mammals, 2 basic helix-loop-helix transcription factors, Circadian Locomotor Output Cycles Kaput (CLOCK) and Brain and Muscle Aryl Hydrocarbon Receptor Nuclear Translocator-Like Protein 1 (ARNTL also named BMAL1), heterodimerize and subsequently bind to conserved E-box sequences in target gene promoters. In this manner, this complex activates the transcription of mammalian Period (Per1, Per2, Per3) and Cryptochrome (Cry1, Cry2) genes.<sup>6</sup> The PERs and CRYs proteins are expressed, post-translationally modified, feedback to inhibit their own transcription and are then degraded to lead to a

<sup>†</sup>These authors contributed equally to this work.

CONTACT Anne-Marie François-Bellan anne-marie.françois@univ-amu.fr Aix Marseille Université, CNRS, CRN2M-UMR7286, Faculté de Médecine Nord, 51 Bd Pierre Dramard CS 80011, 13344 Marseille cedex 15, France.

Color versions of one or more of the figures in this article can be found online at www.tandfonline.com/kncl.

Extra View to: Manon Torres, Denis Becquet, Marie-Pierre Blanchard, Séverine Guillen, Bénédicte Boyer, Mathias Moreno, Jean-Louis Franc and Anne-Marie François-Bellan. Circadian RNA expression elicited by 3'-UTR IRAlu-paraspeckle associated elements. eLife 2016;5:e14837; https://doi.org/10.7554/eLife.14837.001 © 2017 Taylor & Francis



**Figure 1.** Molecular mechanisms of the circadian clockwork in mammals. In mammals, rhythmically transcribed BMAL1 heterodimerizes with CLOCK, and together they bind to target E-boxes in the promoters of Per, Cry, Rev-erb, and Ror. PER and CRY proteins are synthesized in the cytoplasm and may be phosphorylated by CK1 kinases including CK1ɛ/ð. PER/CRY heterodimers translocate to the nucleus where they repress their own transcription, generating a near 24-h feedback loop. A second feedback loop represses or activates the transcription of Bmal1 through the actions of REV-ERB or ROR, respectively. BMAL1/CLOCK heterodimer drives oscillating expression of clock-controlled genes with E-box or RORE containing promoters. Post-transcriptional events in the life cycle of a (pre-) mRNA that have been reported to directly influence the circadian clock and/or to be controlled in a circadian manner include splicing, nuclear retention and cytoplasmic export, regulation by miRNA and polyadenylation at the 3' end. Translation and degradation of the mature mRNA are rhythmic processes as well.

new round of BMAL1:CLOCK mediated transcription (For a review see ref.<sup>7</sup>). Another regulatory loop is mediated by the orphan nuclear receptors, the Retinoic Acid Receptor-Related Orphan Receptor  $\alpha/\beta/\gamma$  (*ROR*  $\alpha/\beta/\gamma$ ) and the Reverse Erb  $\alpha/\beta$  (*Rev-erb*  $\alpha/\beta$ ), that activate and inhibit, respectively, transcription of *Bmal1* through the retinoic acid Receptor Response Element (RRE) in its promoter, leading it to oscillate in a circadian manner (Fig. 1).

In addition to the core regulation at the level of transcription or translation, circadian clock proteins are also subjected to extensive post-translational modifications that appear to control their cellular localization, protein stability, and activity. For example, Casein Kinase I $\epsilon$  and I $\delta$  (CKI $\epsilon/\delta$ ) are known to be critical factors that regulate the turnover of PERs and CRYs in mammals;<sup>8-10</sup> however, kinase CKI $\epsilon$  also activates BMAL1-mediated transcription<sup>9</sup> (Fig. 1).

Importantly, circadian transcription factors not only regulate their own transcription but also regulate the expression of numerous other *clock-controlled genes*<sup>6</sup> [CCGs; (Fig. 1)]. Over the past decade, clock gene transcriptional regulation has been described in many species and tissues, where it drives rhythmic mRNA expression. By use of techniques such as microarrays,<sup>11-13</sup> a large fraction of the mRNA population (up to 10-15% of all mRNAs in a single mammalian tissue<sup>14</sup>) has been shown to display a rhythmic expression that has been initially assumed to result from temporal changes in transcription. However, data from mouse liver demonstrate poor correlation between the activation of a promoter and the amount of the corresponding transcript for genes that are rhythmic at the steady-state level.<sup>15</sup> Actually, with the development of high-throughput sequencing, results obtained in the last years indicate that approximately 43% of the mammalian genome is rhythmic and analysis of circadian nascent RNA has allowed to show that less than 30 % of circadian mRNA are regulated by de novo transcription, suggesting that post-transcriptional regulation contributes mostly to rhythmic mRNA expression (Fig. 1).<sup>15-19</sup> Much of what we initially knew about post-transcriptional regulation came from studies of fungi, plants and flies (For a review see<sup>20</sup>), but circadian post-transcriptional mechanisms involved in rhythmic control of mRNA expression have now also been reported in mammals at many different levels (For a review see ref. 21), such as RNA splicing, poly-adenylation, mRNA stability, mRNA cytoplasmic export and RNA nuclear retention (Fig. 1).

RNAs can be retained in the nucleus by particular bodies called paraspeckles. These nuclear bodies are found in almost all of the cultured cell lines and primary cultures from tissues,<sup>22</sup> except for embryonic stem cells.<sup>23</sup> Paraspeckles are detected as discrete dots found in inter-chromatin space, close to nuclear speckles.<sup>22</sup> A long noncoding RNA, nuclear-enriched abundant transcript one (Neat1) is the structural component (Fig. 2).<sup>23-26</sup> While a short and a long transcript previously identified as MEN<sub>E</sub> (Neat1-1) and MEN $\beta$  (Neat1-2), respectively.<sup>25,27</sup> are generated from the same promoter, Neat1-1 alone cannot induce paraspeckle formation since specific depletion of Neat1-2 leads to disruption of paraspeckles.<sup>25</sup> While paraspeckles detected by RNA FISH of Neat1 appeared as round foci when visualized under a confocal microscope, we showed that they appeared more likely as oblong structures with smaller dimensions after use of a combination of Neat1 RNA FISH and Super Resolution STORM analysis (as designed in Fig. 2). Paraspeckles have been shown to retain in the nucleus RNAs containing duplex structures.<sup>23</sup> This

has been shown for the mouse cationic amino acid transporter 2 (Cat2) transcribed nuclear RNA, Ctn-RNA, an alternatively spliced form of the Cat2 mRNA, which contains a dsRNA structure resulting from inverted short inter-spersed nuclear elements (SINEs) in its 3'-UTR.<sup>28</sup> In primate cells, the most common inverted repeated SINEs are Alu elements. Alu elements are unique to primates and account for almost all of the human SINEs and for more than 10% of the genome and inverted repeat structures (inverted repeated Alu elements [IRAlus]) occur frequently in gene regions.<sup>29</sup> Paraspeckles have been shown to retain in the nucleus mRNAs containing IRAlus in their 3'-UTRs like Nicolin 1 (NICN1) or Lin 28.<sup>23,29</sup>

The nuclear retention of mRNAs containing IRAlus by paraspeckles can be regulated by a methyl-transferase, CARM1 (coactivator-associated arginine methyltransferase1), which control both the binding capacity of some paraspeckle proteins to mRNAs containing IRAlus as well as NEAT1 transcription and then paraspeckle formation.<sup>30</sup> Among the 40 paraspeckle proteins identified thus far,<sup>31</sup> researchers classified 4 RNA-binding proteins, including 3 members of the Drosophila melanogaster behavior human splicing (DBHS) family proteins (NONO, PSPC1 and SFPQ) and RNA-binding motif protein 14 (RBM14) as major paraspeckle protein components (Fig. 2).<sup>22,28,32</sup> We had previously reported that 2 of these major protein components of paraspeckles, namely NONO and SFPQ, display a circadian expression pattern in primary cultures of pituitary cells as well as in a rat pituitary cell line, the GH4C1 cells.<sup>33,34</sup> In this latter cell line, we found thereafter that 2 other major paraspeckle proteins, PSPC1 and RBM14, display also a circadian pattern. All 4 proteins further bind rhythmically to Neat1 and Neat1 itself displays a circadian expression pattern. The expression of the long form of Neat1 RNA, Neat1-2, that is known to be sufficient for the formation of paraspeckle displays also a circadian pattern. In addition we showed that circadian expression of these different components leads to rhythmic variations in paraspeckle number within the cells.<sup>35</sup> Thanks to their circadian expression pattern and given their presumed functions in gene expression through corresponding mRNA nuclear retention, we asked whether paraspeckle bodies can rhythmically retain RNAs in the nucleus leading to a rhythmic expression of the corresponding gene. This hypothesis was first tested with a reporter gene. Indeed, by using



**Figure 2.** Structure and functional implication of paraspeckles in the rhythmic expression of mRNA. Paraspeckles are schematically drawn as oblong structures organized around the short (Neat1–1) and long (Neat1–2) transcripts of the long non-coding Neat1 RNA. Major protein entities of paraspeckles are also shown in the scheme. The roles of paraspeckles in circadian gene expression is schematically represented. The rhythmic number of paraspeckles inside the cells drives a circadian nuclear retention of paraspeckles-associated mRNA and thus leads to a circadian gene expression.

a construct of the EGFP reporter gene fused to an IRAlu and by transfecting the construct into GH4C1 cells, we obtained evidence that IRAlu elements inserted in 3'-UTR of egfp reporter mRNA allow for its circadian retention within the nucleus This rhythmic nuclear retention is abolished after disruption of paraspeckles by siRNA or Neat1 antisens oligonucleotides. Using real-time video microscopy, IRAlu elements inserted in 3'-UTR of egfp reporter mRNA was shown to cause rhythmic cytoplasmic expression of the EGFP protein. Paraspeckles through their circadian expression, could then control circadian expression pattern of a reporter gene containing IRAlu elements in 3'-UTR.35 We then asked whether this circadian post-transcriptional regulation exerted by paraspeckle bodies applies to endogenous genes. To address this issue, it was necessary to determine which mRNA are associated with paraspeckle bodies in the nucleus of GH4C1 cells. To this end, we developed a hybridization-based strategy that uses complementary oligonucleotides to purify Neat1 RNA together with its RNA targets from reversibly cross-linked extracts. Two antisense biotinylated oligonucleotide probes that target accessible regions of Neat1 RNA, as predicted by modeling its secondary structure by bioinformatics, were designed and used for Neat1 RNA specific pull-down whereas one biotinylated irrelevant probe was used for Neat1 RNA non-specific pulldown. Genes which were considered specifically associated with paraspeckles exhibited values of fragment per kilobase per million of mapped reads (FPKM) higher than 1 and were common to the lists obtained

with the 2 specific probes. Overlapping list represented 65% of the list obtained with one probe and 83% of the list obtained with the other one. By comparing this gene list to available data in the literature such as the list of rhythmic mRNA in the mouse pituitary<sup>13</sup> or the mRNA whose rhythm is post-transcriptionally controlled in the liver of mice,<sup>15</sup> we achieved significant recovery rates of 18 and 27%, respectively. These rates are probably underestimated since they come from the comparison of gene lists obtained in different structures and species.

By selecting a few genes from our list, we showed that the mRNA of these genes exhibit a circadian rhythm of their nuclear retention and that this rhythm is abolished when paraspeckles are disrupted.<sup>35</sup> These results obtained on endogenous mRNA, together with results obtained with a reporter gene, allow to conclude that paraspeckles participate in the post-transcriptional control of rhythmic gene expression (Fig. 2). Besides deciphering a post-transcriptional mechanism involved in the circadian regulation of gene expression, a problematic anchored in the field of chronobiology, our study from an endocrine point of view allows to assign a role in the physiology of the pituitary gland to the long non-coding RNA Neat1 whose functions remain enigmatic.

Although it has been shown that paraspeckles can retain in the nucleus genes which have sequences of IRAlu type in their 3'-UTR,<sup>23,28</sup> endogenous mRNA we found associated with Neat1, actually do not contain such sequences or equivalent IR-SINE sequences in our model of rat pituitary cells. This suggests either that paraspeckles can recognize IR-sequences localized anywhere along the mRNA or that doublestranded RNA structures with such IR-sequences are not the only one recognized by paraspeckles. The challenge will be now to elicite the nature of the sequences together with their position along the mRNAs that are involved in the binding of paraspeckles and therefore form the basis of circadian expression.

In conclusion, circadian regulation has been investigated mainly at the transcriptional or the posttranslational level, and RNA-based mechanisms contributing to this regulation are only beginning to emerge. Therefore, the impact of such mechanisms on circadian biology remains to be evaluated. However, given the finding that the majority of circadian mRNAs is not regulated by de novo transcription and given from the study presented here, the high number of mRNA that are potentially regulated by new post-transcriptional mechanism the we described, we predict all these post-transcriptional mechanisms will be shown to play extensive and widespread roles in circadian biology. It is also tempting to imagine that after the identification of mechanisms involved in post-transcriptional control of the circadian clock, connections to diseases will follow. Actually presumably because circadian rhythms play a key role as adaptive mechanisms to permanent changes in the environment, their deregulation is associated with many diseases the best known of which are neuropsychiatric, metabolic, cardiovascular disorders and cancers. A better understanding of the molecular mechanisms that govern the mammalian circadian clock is thus required to first identify connections to diseases, which may then give opportunities for new therapeutic concepts. Therapeutic targeting of paraspeckles could be for instance the basis for new strategies to control circadian rhythms, with the aim of improving human disease associated with rhythmicity dysfunction, in particular rhythmic hormonal dysfunctioning.

#### **Disclosure of potential conflicts of interest**

No potential conflicts of interest were disclosed.

#### Funding

This work was funded by Pfizer-SFE award - Research price Anne-Marie François-Bellan 2014.

#### ORCID

Anne-Marie François-Bellan D http://orcid.org/0000-0002-3278-4642

#### References

- Harmer SL, Panda S, Kay SA. Molecular bases of circadian rhythms. Ann Rev Cell Dev Biol 2001; 17:215-253; PMID:11687489
- [2] Bell-Pedersen D, Cassone VM, Earnest DJ, Golden SS, Hardin PE, Thoma TL, Zoran MJ. Circadian rhythms from multiple oscillators: lessons from diverse organisms. Nat Rev Genet 2005; 6:544-556; PMID:15951747; https:// doi.org/10.1038/nrg1633
- [3] King DP, Takahashi JS. Molecular genetics of circadian rhythms in mammals. Ann Rev Neurosci 2000; 23:713-742; PMID:10845079
- Schibler U, Sassone-Corsi P. A web of circadian pacemakers. Cell 2002; 111:919-922; PMID:12507418; https:// doi.org/12507418
- [5] Hastings MH, Maywood ES, Reddy AB. Two decades of circadian time. J Neuroendocrinol 2008; 20:812-819; PMID:18601704
- [6] Reppert SM, Weaver DR. Molecular analysis of mammalian circadian rhythms. Ann Rev Physiol 2001; 63:647-676; PMID:11181971
- [7] Ko CH, Takahashi JS. Molecular components of the mammalian circadian clock. Hum Mol Genet 2006; 15: R271-R277; https://doi.org/10.1093/hmg/ddl207
- [8] Akashi M, Tsuchiya Y, Yoshino T, Nishida E. Control of intracellular dynamics of mammalian period proteins by casein kinase I epsilon (CKIepsilon) and CKIdelta in cultured cells. Mol Cell Biol 2002; 22:1693-1703; PMID:11865049
- [9] Eide EJ, Vielhaber EL, Hinz WA, Virshup DM. The circadian regulatory proteins BMAL1 and cryptochromes are substrates of casein kinase Iepsilon. J Biol Chem 2002; 277:17248-17254; PMID:11875063; https://doi.org/ 10.1074/jbc.M111466200
- [10] Gallego M, Virshup DM. Post-translational modifications regulate the ticking of the circadian clock. Nat Rev Mol Cell Biol 2007; 8:139-148; PMID:17245414; https:// doi.org/10.1038/nrm2106
- [11] Panda S, Antoch MP, Miller BH, Su AI, Schook AB, Straume M, Schultz PG, Kay SA, Takahashi JS, Hogenesch JB. Coordinated transcription of key pathways in the mouse by the circadian clock. Cell 2002; 109:307-320; PMID:12015981
- [12] Storch K-F, Lipan O, Leykin I, Viswanathan N, Davis FC, Wong WH, Weitz CJ. Extensive and divergent circadian gene expression in liver and heart. Nature 2002; 417:78-83; https://doi.org/10.1038/nature744; PMID:11967526
- [13] Hughes M, Deharo L, Pulivarthy SR, Gu J, Hayes K, Panda S, Hogenesch JB. High-resolution time course analysis of gene expression from pituitary. Cold Spring Harb Symp Quant Biol 2007; 72:381-386;

PMID:18419295; https://doi.org/10.1101/sqb.2007. 72.011

- [14] Lowrey PL, Takahashi JS. Mammalian circadian biology: elucidating genome-wide levels of temporal organization. Annu Rev Genomics Hum Genet 2004; 5:407-441; PMID:15485355; https://doi.org/10.1146/annurev. genom.5.061903.175925
- [15] Menet JS, Rodriguez J, Abruzzi KC, Rosbash M. Nascent-Seq reveals novel features of mouse circadian transcriptional regulation. Elife 2012; 1:e00011; PMID:23150795; https://doi.org/10.7554/eLife.00011
- [16] Partch CL, Green CB, Takahashi JS. Molecular architecture of the mammalian circadian clock. Trends Cell Biol 2013; 24:90-9 PMID:23916625; https://doi.org/10.1016/j. tcb.2013.07.002
- [17] Hurley JM, Loros JJ, Dunlap JC. Circadian oscillators: Around the transcription-translation feedback loop and on to output. Trends Biochem Sci. 2016; 41(10):834-46; PMID:27498225; https://doi.org/10.1016/j.tibs.2016. 07.009
- [18] Koike N, Yoo SH, Huang HC, Kumar V, Lee C, Kim TK, Takahashi JS. Transcriptional architecture and chromatin landscape of the core circadian clock in mammals. Science 2012; 338:349-354; PMID:22936566; https://doi.org/ 10.1126/science.1226339
- [19] Li J, Grant GR, Hogenesch JB, Hughes ME. Considerations for RNA-seq analysis of circadian rhythms. Methods Enzymol 2015; 551:349-367; https://doi.org/10.1016/ bs.mie.2014.10.020
- [20] Kojima S, Shingle DL, Green CB. Post-transcriptional control of circadian rhythms. J Cell Sci 2011; 124:311-320; PMID:21242310; https://doi.org/10.1242/jcs.065771
- [21] Preußner M, Heyd F. Post-transcriptional control of the mammalian circadian clock: implications for health and disease. Pflugers Arch 2016; 468(6):983-91; https://doi. org/10.1007/s00424-016-1820-y
- [22] Bond CS, Fox AH. Paraspeckles: nuclear bodies built on long noncoding RNA. J Cell Biol 2009; 186:637; PMID:19720872
- [23] Chen LL, Carmichael GG. Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA. Mol Cell 2009; 35:467-478; PMID:19716791; https://doi. org/10.1016/j.molcel.2009.06.027
- [24] Clemson CM, Hutchinson JN, Sara SA, Ensminger AW, Fox AH, Chess A, Lawrence JB. An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. Mol Cell 2009; 33:717-726; PMID:19217333; https://doi.org/10.1016/j.molcel.2009. 01.026
- [25] Sasaki YT, Hirose T. How to build a paraspeckle. Genome Biol 2009; 10:227; PMID:19664169; https://doi. org/10.1186/gb-2009-10-7-227

- [26] Sunwoo H, Dinger ME, Wilusz JE, Amaral PP, Mattick JS, Spector DL. MEN epsilon/beta nuclear-retained noncoding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. Genome Res 2009; 19:347-359; PMID:19106332; https:// doi.org/10.1101/gr.087775.108
- [27] Guru SC, Agarwal SK, Manickam P, Olufemi SE, Crabtree JS, Weisemann JM, Kester MB, Kim YS, Wang Y, Emmert-Buck MR. A transcript map for the 2.8-Mb region containing the multiple endocrine neoplasia type 1 locus. Genome Res 1997; 7:725-735; PMID:9253601
- [28] Prasanth KV, Prasanth SG, Xuan Z, Hearn S, Freier SM, Bennett CF, Zhang MQ, Spector DL. Regulating gene expression through RNA nuclear retention. Cell 2005; 123:249-263; PMID:16239143; https://doi.org/10.1016/j. cell.2005.08.033
- [29] Chen LL, Carmichael GG. Gene regulation by SINES and inosines: biological consequences of A-to-I editing of Alu element inverted repeats. Cell Cycle 2008; 7:3294-3301; PMID:18948735; https://doi.org/18948735
- [30] Hu SB, Xiang JF, Li X, Xu Y, Xue W, Huang M, Wong CC, Sagum CA, Bedford MT, et al. Protein arginine methyltransferase CARM1 attenuates the paraspecklemediated nuclear retention of mRNAs containing IRAlus. Genes Dev 2015; 29:630-645; PMID:25792598; https://doi.org/10.1101/gad.257048.114
- [31] Naganuma T, Hirose T. Paraspeckle formation during the biogenesis of long noncoding RNAs. RNA Biol 2013; 10:456-61; PMID:23324609; https://doi.org/ 23324609
- [32] Nakagawa S, Hirose T. Paraspeckle nuclear bodies-useful uselessness? Cell Mol Life Sci 2012; 69:3027-3036; PMID:22476590; https://doi.org/10.1007/s00018-012-0973-x
- [33] Guillaumond F, Boyer B, Becquet D, Guillen S, Kuhn L, Garin J, Belghazi M, Bosler O, Franc JL, François-Bellan AM. Chromatin remodeling as a mechanism for circadian prolactin transcription: rhythmic NONO and SFPQ recruitment to HLTF. FASEB J. 2011; 25:2740-2756; PMID:21507896; https://doi.org/10.1096/fj.10-178616
- Becquet D, Boyer B, Rasolonjanahary R, Brue T, Guillen S, Moreno M, Franc JL, François-Bellan AM. Evidence for an internal and functional circadian clock in rat pituitary cells. Mol Cell Endocrinol 2014; 382:888-898; PMID:24239982; https://doi.org/10.1016/j.mce.2013. 11.004
- [35] Torres M, Becquet D, Blanchard MP, Guillen S, Boyer B, Moreno M, Franc JL, Fran , c ois-Bellan AM. Circadian RNA expression elicited by 3 0 -UTR IRAlu-paraspeckle associated elements. Elife 2016;5:e14837 https://doi.org/ 10.7554/eLife.14837

# Recherche de déterminants moléculaires impliqués dans la rétention d'ARNm par les paraspeckles

# Introduction

L'étude du rôle de Neat1 et des paraspeckles dans la rythmicité circadienne du transcriptome hypophysaire a permis de montrer qu'un grand nombre d'ARNs hypophysaires s'associent à Neat1 (Torres et al 2016). Les données de la littérature suggèrent que la reconnaissance des ARNs par les paraspeckles est conditionnée par la présence de motifs dans ces ARNs qui permettent leur liaison à un des éléments des paraspeckles. Le principal motif décrit à ce jour est le motif IRAlu (Prasanth et al 2005; Chen et al 2008; Mao et al 2011) . Ce motif correspond à deux séquences Alu qui sont des séquences de la famille de courts éléments nucléaires intercalés (ou SINE pour « short interspersed nuclear elements »), qui sont inversées et forment des boucles d'ARN double brin. Ces motifs sont particulièrement nombreux dans la région 3'UTR des ARNs. Ces boucles d'ARN double brin sont reconnues par l'enzyme ADAR responsable de leur édition A->I (Hundley, & Bass 2010; Liscovitch et al 2014) . La structure IRAlu éditée présente dans la région 3' UTR serait alors spécifiquement reconnue par un ou plusieurs des éléments des paraspeckles.

Bien que les séquences Alu soient spécifiques de l'homme et des primates, on trouve chez les rongeurs des séquences SINE équivalentes aux séquences Alu des primates. Etonnamment, la recherche par bioinformatique de motifs IRSINEs dans les régions 3' UTR des ARNs trouvés associés aux paraspeckles n'a pas permis de mettre en évidence la présence de ces motifs dans notre modèle d'étude. Par conséquent, soit les séquences de reconnaissance des ARNs que nous avons trouvés associés à Neat1 sont localisées dans d'autres régions que la région 3'UTR ou bien d'autres motifs présents dans la région 3'UTR de ces ARNs permettent leur reconnaissance par les paraspeckles. Pour aborder cette question et poursuivre le travail présenté précédemment, nous avons réalisé une recherche ciblée sur des ARNs que nous avons sélectionnés, afin de déterminer si la région 3'UTR de ces ARNs est déterminante pour leur interaction avec les paraspeckles et dans ce cas quels sont les motifs présents dans cette région qui permettent leur reconnaissance.

L'étude a été réalisée sur trois ARNs dont nous avons montré la rétention nucléaire par association aux paraspeckles dans notre précédente étude (Torres et al 2016). Ces ARNm sont d'une part ceux de l'hormone de croissance Gh et de l'hormone Prolactine (Prl), qui sont les deux hormones exprimées dans la lignée hypophysaire somatolactotrope GH4C1 et d'autre part ceux de la calréticuline (Calr), dont plusieurs expériences réalisées au sein de l'équipe ont montré une association spécifique avec les paraspeckles ainsi qu'une régulation circadienne exclusivement dépendante des paraspeckles. Pour ces trois ARNm nous avons testé dans un premier temps l'implication de la région 3'UTR dans l'association aux paraspeckles. Puis, par une approche de fragmentation de la région 3'UTR de ces ARNm nous avons tenté de préciser la région contenant le ou les motifs de reconnaissance par les paraspeckles.

# Matériel et méthodes

# **Cellules GH4C1**

Les cellules GH4C1 sont cultivées en monocouche dans une étuve à 37°C, saturée en H20 et en présence de 5 % de CO2. Le milieu de culture (HamF10) est enrichi avec 15 % de sérum de cheval (HS) et 2 % de sérum de veau fœtal (SVF) et contient 1 % d'antibiotiques (0,5 % de streptomycine et 0,5 % de pénicilline).

# **Clonage des séquences 3'UTR**

Les séquences 3'UTR des ARNs de Gh, Prl et Calr on été clonées à partir des ARNs totaux des cellules GH4C1. Les ARNs totaux de GH4C1 (5µg) ont été transformés en ADNc par reverse transcription (RT) avec le kit SuperScript<sup>™</sup> Reverse Transcriptase (ThermoFisher).

A partir de l'ADNc, les régions 3'UTR de Gh, Prl et Calr ont été amplifiées par PCR avec la Phusion Green Hot Start II High-Fidelity DNA Polymerase (ThermoFisher) et des primers situés au début et à la fin de la région 3'UTR (table 1) (Scotto-Lavino et al 2006). Deux étapes successives permettent d'assurer la spécificité de l'amplification.

Les séquences ainsi obtenues ont été insérées dans le vecteur plasmidique d'expression de la protéine EGFP (pEGFP) (Clontech, Mountain View, CA). Le vecteur pEGFP a été ouvert au niveau du site de restriction de Sma1, et les séquences 3'UTR ont été insérées par ligation avec la T4 Polynucleotide Kinase (New England BioLabs inc., NEB).

Les plasmides contenant les différentes constructions ont été amplifiés à l'aide de bactéries « Subcloning EfficiencyTM DH5αTM Competent Cells » (Life Technologies) transformées par choc thermique. Les bactéries ont ensuite été ensemencées dans des boites d'agar contenant de la Kanamycine. L'insertion des séquences 3'UTR dans les plasmides a été vérifiée par PCR puis par séquençage (Beckman Coulter Genomics) pour confirmer l'exactitude de la séquence insérée. Enfin, les bactéries contenant les bonnes constructions ont été amplifiées pour pouvoir purifier en grande quantité le plasmide avec le kit « NucleoBond Xtra Midi » (Macherey-Nagel).

# Création de lignées stables GH4C1 exprimant les constructions pegfp-3'UTR

Les cellules GH4C1 ont été transfectées avec les constructions plasmidiques exprimant le gène de résistance à la néomycine avec l'agent transfectant Lipo3000 (Invitrogen, Cergy Pontoise, F). Deux jours après la transfection, la sélection des cellules positives pour la transfection est réalisée pendant plusieurs semaines avec 250 mg/ml de G418 (Invitrogen). Après sélection, l'expression de la construction est vérifiée par amplification de la séquence egfp par RT-qPCR.

# RNA pull-down de Neat1

Les lignées GH4C1 exprimant les différentes constructions egfp ont été fixées au paraformaldéhyde 1% afin de stabiliser les complexes formés par Neat1 et ses partenaires ARNs et protéiques. Les cellules ainsi fixées sont lysées. Une étape de traitement aux ultrasons permet de fragmenter l'ADN et de fluidifier les échantillons. L'isolement de Neat1 et de ses partenaires est obtenu par hybridation avec des sondes oligonucléotidiques anti-sens vis-à-vis de Neat1. Des sondes oligonucléotidiques biotinylées non spécifiques de Neat1 permettent de déterminer la liaison non-spécifique. Les ARNs sont ensuite purifiés par des étapes de rinçage et de dégradation de l'ADN et des protéines. Le protocole détaillé de cette technique de RNA pull-down fait l'objet de notre publication dans le Journal of Visualized Experiments {Torres 2018, sous presse}.

# Séparation des ARNs nucléaires et cytoplasmiques

Pour évaluer la rétention nucléaire de l'ARN egfp dans les lignées stables exprimant les différentes constructions, les ARNs ont été quantifiés dans la fraction nucléaire et la fraction cytoplasmique. Pour séparer les compartiments nucléaires et cytoplasmiques les cellules GH4C1 exprimant les différentes constructions sont incubées 5min dans un tampon de lyse ménagée permettant de lyser la membrane cytoplasmique sans lyser la membrane nucléaire (10 mmol/L de Tris pH 7.4, 3 mmol/L de MgCl2, 10 mmol/L de NaCl, et 0.5% de NP-40). La fraction cytoplasmique est obtenue par centrifugation (510 rcf 5 min à 4°C). La préparation nucléaire est lavée deux fois par le tampon de lyse

ménagée. Les ARNs sont extraits à l'aide du kit RNA II ou RNA XS (Macherey Nagel, Hoerdt, France) dans chacune des deux fractions.

# Analyse de l'expression des ARN

Les ARNs purifiés (500ng) sont ensuite convertis en ADNc par Reverse Transcription avec le kit RNA to cDNA (Applied Biosystem, Courtaboeuf, France). La mesure relative du niveau d'expression des ARNs est faite par qPCR en temps réel avec le mix Fast SYBER Green (Applied Biosystems) sur l'appareil 7500 fast Real-Time qPCR system (Applied Biosystems). Les primers utilisés sont listés dans la Table 2. L'accumulation des ARNs est normalisée relativement au niveau de l'ARNm de Gapdh. Les résultats sont ensuite exprimés en valeurs relatives par rapport aux valeurs obtenues dans les échantillons non spécifiques, puis par rapport à la construction de référence IRAluegfp, en utilisant la méthode de calcul du  $\Delta\Delta$ Ct décrite par M.W.Pfaffl (Pfaffl 2001).

# Analyse statistique

La significativité des différences entre les différents groupes comparés a été déterminée par analyse simple ANOVA suivi d'un test-t. Les valeurs sont considérées significativement différentes pour p<0.05 (\*), p<0.01 (\*\*), p<0.0001(\*\*\*\*).
#### **Résultats :**

# Association de la région 3'UTR des ARNm Gh, Prl et Calr à Neat1, mise en évidence par RNA pull-down

Pour déterminer si la région de reconnaissance par les paraspeckles des ARNm de Gh, Prl et Calr est située dans la région 3'UTR, les séquences 3'UTR de ces trois ARNm ont été clonées et insérées en aval de la région codante et avant le signal de polyadénylation du gène codant pour la protéine EGFP. Ces constructions ont été transfectées de manière stable dans la lignée cellulaire GH4C1. Après RNA pull-down avec des sondes spécifiques dirigées contre Neat1 et des sondes non-spécifiques, la mesure du niveau d'enrichissement en gène rapporteur egfp avec les sondes spécifiques (OS) comparées aux sondes non-spécifiques (ONS), permet d'évaluer l'association des différentes constructions à Neat1. L'enrichissement en gène rapporteur obtenu avec les différentes constructions est rapporté à celui obtenu avec la construction de référence c'est-à-dire celle exprimant le motif IRAlu en aval du gène rapporteur egfp (IRAlu-egfp). Bien que quelque soit la construction on observe que l'ARN de l'egfp est significativement enrichi après utilisation de sondes spécifiques anti-sens vis-à-vis de Neat1 comparée à des sondes non-spécifiques (Fig9.A), le niveau d'enrichissement obtenu avec les régions 3'UTR des 3 ARNm est significativement moindre comparé à l'enrichissement obtenu avec la construction IRAlu-egfp (Fig9.B).

#### Rétention nucléaire de la région 3'UTR des ARNm Gh, Prl et Calr

La répartition entre la fraction nucléaire et la fraction cytoplasmique des ARNm du gène rapporteur egfp permet d'évaluer l'influence de la région 3'UTR de chacun des 3 ARNs sur la rétention nucléaire de la construction correspondante. Le ratio entre l'expression nucléaire et l'expression cytoplasmique de l'egfp de chacune des constructions a été exprimé relativement au ratio obtenu avec la construction IRAluegfp. Les régions 3'UTR de Gh et de Calr entrainent une rétention nucléaire de l'egfp entraînée par la séquence IRAlu (Fig9.C). En revanche, la rétention nucléaire de l'egfp entraînée par la 3'UTR de la Prl est significativement moindre comparée à celle induite par la séquence IRAlu (Fig9.C).

Il apparaît donc que les régions 3'UTR de Gh et Calr induisent une association à Neat1 plus faible que celle du motif IRAlu et cependant leur rétention nucléaire est tout aussi forte. La région 3'UTR de Prl en revanche n'est que faiblement retenue dans le noyau malgré une association à Neat1 égale à celle des séquences 3'UTR de Gh et Calr.

## Recherche dans les régions 3'UTR des ARNm de Gh et Calr des sous-régions impliquées dans l'association à Neat1

Pour préciser les régions au sein des queues 3'UTR qui contiennent le ou les motifs de reconnaissance par les paraspeckles, une approche de fragmentation a été réalisée. De nouvelles constructions du gène rapporteur egfp contenant différents fragments des séquences 3'UTR de Gh et de Calr (Fig10.A) ont été transfectées de manière stable dans la lignée GH4C1.

L'ARN de l'egfp est significativement enrichi après utilisation de sondes spécifiques anti-sens vis-à-vis de Neat1 comparée à des sondes non-spécifiques (Fig10.B). Les fragments issus de la région 3'UTR de l'ARN de la Gh, montrent que le fragment 2 (egfp-3'GH2) entraine un enrichissement en gène rapporteur équivalent à celui obtenu avec la séquence de référence IRAlu-egfp alors que le fragment 1 (egfp-3'GH1) entraine un enrichissement moindre (Fig10.C). Les fragments de la région 3'UTR de l'ARNm de la Calr montrent que le fragment 3 (egfp-3'CR) entraine un enrichissement en gène rapporteur équivalent à celui obtenu avec la séquence de référence IRAlu-egfp alors que les fragment 1 et 2 (egfp-3'CR1 et egfp-3'CR2) entrainent un enrichissement moindre (Fig10.C).

Les fragments egfp-3'GH2 et egfp-3'CR3 ont donc une plus forte affinité pour Neat1 que les autres fragments, et cette affinité est équivalente à celle du motif IRAlu.

#### Rétention nucléaire des fragments de la région 3'UTR des ARNm Gh et Calr

Ici encore, le ratio entre l'expression nucléaire et l'expression cytoplasmique de l'egfp de chacune des constructions a été exprimé relativement au ratio obtenu avec la construction IRAlu-egfp. Bien que les fragments de la région 3'UTR de la Gh entrainent tous deux une rétention nucléaire de l'egfp significativement moindre comparée à celle induite par la séquence IRAlu, le fragment 1 (egfp-3'GH1) entraine une rétention du gène rapporteur plus importante que le fragment 2 (egfp-3'GH2) (Fig10.D).

Pour ce qui concerne les fragments de la région 3'UTR de la Calr, bien que les 3 fragments soient moins efficaces que la séquence IRAlu pour retenir le gène rapporteur dans le noyau, les fragments 1 et 3 (egfp-3'CR1, egfp-3'CR3) ont une efficacité comparable et supérieure au fragment 2 (egfp-3'CR2) (Fig10.D).

Il apparaît au vu de ces résultats que la liaison à Neat1 des sous-régions egfp-3'GH2 et egfp-3'CR3, bien qu'équivalente à celle d'un motif IRAlu, n'est pas suffisante pour induire le même niveau de rétention nucléaire de l'ARN egfp obtenu avec ce motif.



## Figure 9 : Rôle des régions 3'UTR des ARNs de Gh, Prl et Calr dans l'association à Neat1 et la rétention nucléaire de l'ARNm de l'egfp

A- L'association à Neat1 déterminée après Neat1 RNA pull-down est quantifiée par mesure de l'enrichissement en ARNm egfp. Pour les sondes spécifiques (OS) et non spécifiques (ONS), l'enrichissement est exprimé relativement à la valeur obtenue dans les inputs. (\*\*)p<0,01 (\*)p<0.05 significativement différent entre les sondes OS et ONS.

B- L'association à Neat1 déterminée après Neat1 RNA pull-down est quantifiée par mesure de l'enrichissement en ARNm egfp exprimé relativement à l'enrichissement obtenu avec la construction egfp contenant une séquence IRAlu (IRAlu-egfp).(\*\*\*\*)p<0,0001 significativement différent de la construction IRAlu-egfp.

C- Ratio de l'expression nucléaire versus cytoplasmique de l'ARNm de l'egfp des différentes constructions. Le ratio obtenu pour chaque construction est rapporté à celui obtenu avec la construction contenant la séquence de référence IRAlu-egfp . (\*) p<0.05 significativement différent de la construction IRAlu-egfp.



Figure 10

## Figure 10: Rôle de différents fragments des régions 3'UTR des ARNs de Gh et Calr dans l'association à Neat1 et la rétention nucléaire de l'ARNm de l'egfp

A- Schéma des fragments clonés à partir des régions 3'UTR de l'ARNm de Gh et de la Calr. B- L'association à Neat1 déterminée après Neat1 RNA pull-down est quantifiée par mesure de l'enrichissement en ARNm egfp. Pour les sondes spécifiques (OS) et non spécifiques (ONS), l'enrichissement est exprimé relativement à la valeur obtenue dans les inputs. (\*\*\*) p<0,001 (\*\*)p<0,01 (\*)p<0.05 significativement différent entre les sondes OS et ONS.

C- L'association à Neat1 déterminée après Neat1 RNA pull-down est quantifiée par mesure de l'enrichissement en ARNm egfp exprimé relativement à l'enrichissement obtenu avec la construction egfp contenant une séquence IRAlu (IRAli-egfp).

Sur le diagramme de gauche, l'enrichissement relatif avec les constructions egfp contenant les deux fragments de la région 3'UTR de Gh. Sur le diagramme de droite, l'enrichissement relatif avec les constructions egfp contenant les trois fragments de la région 3'UTR de la Calr. (\*) p<0.05 (\*\*) p<0.01 significativement différent de la construction IRAlu-egfp.

D- Ratio de l'expression nucléaire versus cytoplasmique de l'ARNm de l'egfp avec les différentes constructions. Le ratio obtenu pour chaque construction est rapporté à celui obtenu avec la construction contenant la séquence de référence IRAlu-egfp.

. (\*\*\*\*) p<0.0001 significativement différent de la construction IRAlu-egfp.

(\$\$) p<0.01 significativement différent entre les constructions.

#### Discussion

Parmi les gènes retrouvés associés à Neat1 dans nos expériences de Neat1 RNA pulldown suivi d'un séquençage haut débit des ARNs, se trouvent les deux gènes codant pour les hormones spécifiques de notre modèle cellulaire, l'hormone de croissance et la prolactine. Pour aborder la caractérisation des éléments présents dans les ARNs qui déterminent leur capacité à lier les paraspeckles, ces deux gènes hormonaux ont été sélectionnés. Un autre gène, celui de la calréticuline, a été également sélectionné. En effet, bien que ce gène soit ubiquitaire, des données obtenues dans l'équipe montrent que ce gène est fortement associé aux paraspeckles. Ces trois ARNs ont donc été sélectionnés pour déterminer en premier lieu si la région 3'UTR de ces ARNs est responsable de leur association aux paraspeckles.

Par RNA pull-down nous avons pu montrer que les régions 3'UTR de Gh, Prl et Calr s'associent de manière spécifique à Neat1. Cependant l'efficacité des régions 3'UTR de ces ARNm à entrainer une association du gène rapporteur egfp à Neat1 est moindre comparée à l'efficacité de la séquence IRAlu considérée comme la séquence de référence. En dépit de cette association moindre, les régions 3'UTR de Gh et Calr sont capables de retenir dans le noyau l'ARNm du gène rapporteur aussi efficacement qu'une séquence IRAlu. Il est donc possible que des mécanismes autres que l'association aux paraspeckles, concourent à la rétention nucléaire du gène rapporteur induite par les régions 3'UTR de Gh et de Calr. La région 3'UTR de la Prl quant à elle, est non seulement moins efficace qu'une séquence IRAlu pour entrainer l'association du gène rapporteur egfp à Neat1 mais également moins efficace pour entraîner sa rétention dans le noyau. Dès lors, nous avons postulé que la région 3'UTR de l'ARNm de la Prl ne participe que partiellement à l'association de l'ARNm de la Prl aux paraspeckles et que d'autres régions de l'ARN pouvaient être impliquées. Ceci nous a conduits à éliminer l'ARNm de la Prl dans la suite de notre étude.

Afin de préciser au sein des régions 3'UTR de Gh et de Calr, les sous-régions impliquées dans l'interaction avec les paraspeckles, nous les avons fragmentées respectivement en deux et trois fragments. Après Neat1 RNA pull-down il apparaît que tous les fragments générés permettent d'entraîner une liaison spécifique du gène rapporteur à

Neat1 ce qui semble indiquer que l'association à Neat1 peut se faire par plusieurs sousrégions de la région 3'UTR d'un ARN, ce qui dès lors pourrait impliquer plusieurs motifs de reconnaissance. Cependant les différents fragments générés ne s'associent pas avec la même efficacité à Neat1, ce qui semble indiquer que certaines sous-régions participent préférentiellement à la reconnaissance de l'ARNm par les paraspeckles. Il est donc tentant de postuler que l'efficacité de liaison à Neat1 pourrait dépendre du nombre de domaines d'interaction présents dans la région 3'UTR d'un ARNm.

Chacun des fragments générés pris isolément présente cependant une efficacité réduite à retenir le gène rapporteur egfp dans le noyau, comparée à l'efficacité de la région 3'UTR entière. On peut par conséquent supposer que l'efficacité de la région 3'UTR entière, aussi bien pour l'ARN de la Gh que pour celui de la Calr, nécessite l'ensemble des sous-régions, en particulier puisque, comme discuté dans le paragraphe précédent, chacune de ces sous-régions pourrait contenir des domaines d'interaction avec Neat1. On ne peut toutefois pas exclure que la fragmentation des régions 3'UTR ait pu induire une modification importante de la structure secondaire des ARNs qui pourrait être déterminante pour leur reconnaissance par les paraspeckles. En effet, la reconnaissance d'un ARNm par les paraspeckles pourrait être conditionnée non seulement par la séquence primaire mais aussi par la configuration spatiale de sa région 3'UTR.

A partir des trois ARNm sélectionnés, notre étude nous permet de suggérer que la participation de la région 3'UTR d'un ARNm à son association aux paraspeckles varie en fonction de l'ARNm considéré. En effet, on a vu que pour l'ARNm de la Prl, des régions autres que la région 3'UTR pourraient participer à cette association. A cet égard il est intéressant de noter que l'hétérodimèrisation de deux des protéines du paraspeckles, NONO et SFPQ, permet la formation d'un site de liaison ayant une forte affinité pour la région 5'UTR des ARNs (Duvignaud et al 2016). Il reste néanmoins à déterminer si la liaison de ce dimère protéique aux régions 5'UTR implique les paraspeckles et entraîne une rétention nucléaire des ARNm concernés.

Pour les ARNm Gh et Calr, il est apparu au travers de notre étude que plusieurs sousdomaines de la région 3'UTR pourraient être collectivement impliqués dans l'association aux paraspeckles. Dans ce cas, la rétention nucléaire par les paraspeckles serait moins la résultante d'une reconnaissance par un site spécifique que la sommation des influences de plusieurs sites d'interaction. Ces sites d'interaction pourraient correspondre aux différents motifs de liaison des nombreuses protéines dont sont composés les paraspeckles et qui sont majoritairement des protéines de liaison aux ARNs (Naganuma et al 2012).

Par ailleurs, nous avons montré que les régions 3'UTR entières de Gh, Prl et Calr malgré une faible association à Neat1 (en comparaison à la séquence IRAlu) entraînent une forte rétention nucléaire du gène rapporteur. Il a récemment été montré que certains ARNs retenus par les paraspeckles pouvaient aussi être retenus par d'autres éléments du noyau. Ainsi la rétention nucléaire de l'ARN de CTN (mouse cationic amino acid transporter 2 (mCAT2) transcribed nuclear RNA) bien qu'impliquant les paraspeckles (Prasanth et al 2005) n'est pas complètement abolie après inhibition de la formation des paraspeckles ce qui démontre que d'autres structures nucléaires permettent la rétention de cet ARN (Anantharaman et al 2016). Dans le cas des ARNs de la Gh, la Prl et la Calr, il est donc possible que la rétention nucléaire soit un processus robuste assuré par plusieurs structures. Ces mécanismes pourraient viser à réguler plus finement l'expression de ces trois gènes.

Il apparaît donc que les mécanismes permettant l'association des ARNs à la structure des paraspeckles sont divers et complexes. Ils pourraient non seulement impliquer différentes régions le long de la séquence des ARNs en particulier les régions 3'UTR mais aussi les régions 5'UTR ainsi que plusieurs sous-domaines différents à l'intérieur d'une même région. Du fait de cette complexité, une approche bio-informatique permettant la recherche de motifs communs ou de structures secondaires communes aux ARNs associés aux paraspeckles sera envisagée. **Discussion et Perspectives** 

## 1 – Neat1 RNA pull-down : mise au point et apport de la technique

#### Apport du RNA pull-down pour l'étude des IncRNAs

Les lncRNAs malgré leur nombre important chez les mammifères représentent une catégorie d'ARNs non codants dont les mécanismes et les fonctions sont encore peu décrits à ce jour. L'un des rôles majeurs des lncRNAs apparaît être la régulation de l'expression des gènes. Cette régulation se fait à différents niveaux, de la régulation épigénétique en passant par la transcription jusqu'à la maturation des ARNs de diverses catégories. Ainsi Malat1 va cibler des ARNs immatures dans le noyau (Spector, & Lamond 2011), alors que Neat1 au travers des paraspeckles va interagir dans le noyau avec des ARNs matures et que LincRNA-p21 lui va être localisé dans le cytoplasme et réguler en particulier les ARNs matures riches en polysomes (Yoon et al 2012). De plus, les lncRNAs sont aussi impliqués dans la régulation de l'expression des ARNs non-codants. Des analyses transcriptomiques ont ainsi permis de mettre en évidence d'importantes interactions entre les lncRNAs et les micro-ARNs (miRNAs) (Jalali et al 2013). Le développement de techniques permettant l'étude des interactions entre des lncRNAs et des ARNs constitue donc un enjeu majeur dans la compréhension du rôle de ces lncRNAs dans le fonctionnement cellulaire.

Il existe des stratégies permettant d'étudier les interactions entre un lncRNA et ses partenaires protéiques comme les approches morphologiques couplant la détection par hybridation in situ du lncRNA et la détection par immunocytochimie de la protéine associée ou encore les techniques d'immunoprécipitation des ARNs. Les protéines partenaires étant des protéines de liaison aux ARNs, les ARNs que l'on identifie comme des cibles de ces protéines sont donc des cibles potentielles du lncRNA dont ces protéines sont partenaires. Cependant ce type d'approche qui est indirect est restreint aux ARNs associées à un lncRNA par l'intermédiaire de protéines partenaires et peut par ailleurs donner lieu à des résultats erronés quand les protéines partenaires ne sont pas exclusivement associées au lncRNA. C'est le cas des différentes protéines des paraspeckles que l'on retrouve également en dehors des paraspeckles dans la cellule. De ce fait, dans le cas des paraspeckles pour déterminer directement les cibles ARNs des paraspeckles, il faut pouvoir isoler le lncRNA Neat1 lui même et non pas les protéines des paraspeckles. Cependant, au moment où nous avons entrepris notre étude, les techniques publiées permettaient seulement l'étude des interactions entre un lncRNA et la chromatine. Nous avons donc adapté ces protocoles afin d'obtenir une méthode d'étude des interactions entre un lncRNA et des ARNs. En optimisant les conditions de fixation des interactions moléculaires au sein des cellules, ce sont tous les ARNs associés à un lncRNA qui peuvent ainsi être extraits. En couplant cette méthode à du séquençage haut débit des ARNs, la liste de l'ensemble des ARNs cibles d'un lncRNA peut alors être obtenue.

Le RNA pull-down constitue ainsi l'un des premiers protocoles permettant l'extraction directe d'un lncRNA avec l'intégralité de ses cibles ARNs.

#### Mise au point du protocole de RNA pull-down:

La méthode du RNA pull-down présente deux étapes clés dans sa réalisation : d'une part la sélection de sondes oligonucléotidiques spécifiques du lncRNA d'intérêt et d'autre part la préparation adéquate du matériel cellulaire utilisé. Ces deux étapes doivent être ajustées en fonction des caractéristiques du lncRNA et du modèle cellulaire étudiés.

La conception des sondes oligonucléotidiques est une étape cruciale dont vont dépendre l'affinité et la spécificité des sondes pour le lncRNA d'intérêt. Deux protocoles dont nous nous sommes inspirés pour développer notre propre méthode de RNA pull-down, proposent des techniques différentes pour la modélisation de ces sondes. Ces protocoles Chromatin isolation by RNA purification (ChIRP) et Capture Hybridization Analysis of RNA Targets (CHART) permettent l'extraction d'un lncRNA avec ses sites de fixation à la chromatine. Dans le protocole de ChIRP, un pool de sondes qui s'hybrident sur toute la longueur du lncRNA est utilisé. Dans le protocole de CHART, les propriétés de la RNase-H sont utilisées pour détecter expérimentalement les régions du lncRNA accessibles à l'hybridation. Trois sondes sont ainsi sélectionnées et utilisées en cocktail.

Dans le protocole de ChIRP, la synthèse d'un très grand nombre d'oligonucléotides implique un investissement financier important qui peut être réduit par l'utilisation d'un plus petit nombre de sondes comme dans le protocole de CHART. Pour déterminer les régions accessibles à l'hybridation de manière plus rapide et moins couteuse qu'une approche expérimentale utilisant la RNase-H, la modélisation par bioinformatique de la structure secondaire d'un lncRNA est un outil alternatif. Nous proposons ainsi une stratégie de conception des sondes oligonucléotidiques spécifiques d'un lncRNA basée sur une analyse bio-informatique des régions du lncRNA disponibles pour l'hybridation. De plus, ces sondes ainsi conceptualisées s'étant révélées aussi efficaces individuellement qu'en cocktail, elles peuvent être utilisées séparément ce qui permet par une analyse comparative des résultats obtenus avec ces différentes sondes d'évaluer la spécificité des résultats. Après conception il est toutefois nécessaire de valider les sondes expérimentalement. En effet, les prédictions de structure secondaire d'un lncRNA ne prennent pas en compte les interactions que forment les lncRNAs avec différents partenaires ni la manière dont ces interactions vont influencer la configuration du lncRNA au sein de la cellule. Or ces configurations seront spécifiques d'un lncRNA et modifieront son accessibilité à une hybridation par des sondes d'une manière qui ne peut être prédite par des outils bio-informatiques. Ainsi dans les paraspeckles, le lncRNA Neat1 va être associé à une trentaine de protéines de liaison aux ARNs formant des sous-domaines nucléaires compacts qui rendra difficile l'accès et l'hybridation des sondes oligonucléotidiques sélectionnées (Nakagawa et al 2014; West et al 2016) . Le lncRNA BACE1-AS va lui s'apparier sur toute sa longueur avec l'ARNm BACE1 dont il est complémentaire (Faghihi et al 2008).

La préparation du matériel nucléaire, autre étape clé du protocole de RNA pull-down, doit être adaptée au modèle cellulaire ou tissulaire utilisé. La préparation que nous avons proposée dans notre mise au point permet l'étude des interactions moléculaires aussi bien « in vitro » dans des lignées cellulaires en culture (entre 10<sup>6</sup> et 10<sup>7</sup> cellules) qu' « ex vivo » dans des prélèvements de tissus (entre 1 et 10mg).

Cette préparation du matériel cellulaire nécessite d'une part la fixation (ou Crosslinking) du matériel afin d'assurer la stabilité des interactions entre le lncRNA d'intérêt et ses cibles ARNs et d'autre part le traitement par ultrasons des échantillons qui permet de fragmenter l'ADN qui augmente la viscosité de l'échantillon et empêche sa manipulation. Ces étapes permettent l'accès des sondes oligonucléotidiques au lncRNA et l'extraction du lncRNA auquel restent liés les ARNs. Ceci permet d'isoler aussi bien les ARNs liés directement au lncRNA par une interaction de type ARN/ARN que les ARNs liés aux protéines partenaires du lncRNA. Cette méthode ne permet donc pas de distinguer les ARNs liés de manière directe au lncRNA, de ceux liés par l'intermédiaire des protéines partenaires de ce lncRNA.

Pour poursuivre la caractérisation des mécanismes de rétention des ARNs par les paraspeckles il sera nécessaire de déterminer dans quel type de liaison, ARN/ARN ou protéine/ARN, sont engagés les ARNs associés aux paraspeckles. Pour ce faire, nous mettrons à profit les caractéristiques d'un agent de fixation dérivé du psoralène, le 4'aminomethyltrioxalen (AMT), qui génère la réticulation des bases uridines dans les ARNs sans agir sur les protéines (Engreitz et al 2014). En remplaçant le paraformaldéhyde par le AMT il nous sera alors possible d'extraire par RNA pull-down de Neat1 uniquement les ARNs liant les paraspeckles par des liaisons directes ARN/ARN. Après séquençage haut débit des ARNs, la comparaison de la liste des ARNs ainsi obtenue avec celle des ARNs trouvés associés aux paraspeckles dans notre première étude, nous permettra de distinguer les ARNs associés aux paraspeckles par des liaisons ARN/ARN.

## 2 – Structure des paraspeckles dans les cellules hypophysaires

#### Caractérisation des paraspeckles dans la lignée GH4C1

La formation des paraspeckles est dépendante de l'expression de Neat1 ainsi que de son association à plusieurs protéines de liaison aux ARNs en particulier NONO, SFPQ et RBM14 (Sasaki et al 2009; Naganuma et al 2012). Plusieurs études « in vitro » et « in vivo » montrent que les paraspeckles sont exprimés dans toutes les cellules de mammifères, à l'exception des cellules embryonnaires (Chen, & Carmichael 2009; Nakagawa et al 2011). Cependant le nombre de paraspeckles varie fortement en fonction du type cellulaire étudié. De plus, la quantification respective des deux isoformes de Neat1 met en évidence une grande disparité entre les différents tissus non seulement en terme de niveau d'expression globale mais aussi en terme de proportions respectives des deux isoformes. Il apparaît notamment que le cerveau est un des tissus ayant la plus faible expression des deux isoformes (Nakagawa et al 2011).

Dans les cellules hypophysaires GH4C1, nous avons clairement identifié la présence de paraspeckles non seulement en superposant deux à deux les marquages immunocytochimiques des protéines des paraspeckles mais aussi par identification directe de Neat1 par hybridation in situ de fluorescence. Nous avons de plus montré par immunoprécipitation des ARNs que les protéines des paraspeckles sont bien associées à Neat1.

De manière intéressante, Nakagawa et al. rapportent que l'isoforme Neat1-2, qui est l'isoforme indispensable à la formation des paraspeckles, est particulièrement exprimée dans les tissus spécialisés dans les fonctions sécrétrices (Nakagawa et al 2011). En accord avec ces données, la présence de paraspeckles dans la lignée de cellules hypophysaires GH4C1 mais aussi notre mise en évidence de l'expression de Neat1-2 dans l'hypophyse de souris, pourraient donc être associées à la fonction sécrétrice de la glande hypophysaire.

#### Structure des paraspeckles dans la lignée GH4C1

En utilisant des techniques de microscopie en super-résolution, nous avons aussi précisé la taille et la forme des paraspeckles dans les cellules hypophysaires GH4C1. En analysant les marquages obtenus par hybridation in situ de Neat1 en microscopie confocale les paraspeckles apparaissent de forme ronde avec un diamètre d'environ 360nm. En revanche, après reconstruction en super résolution STORM les paraspeckles sont identifiés comme des structures de plus petite taille et de forme ovale. Ces résultats sont en adéquation avec la caractérisation princeps par microscopie électronique de zones inter-chromatiniennes associées aux granules (IGAZ) identifiées ultérieurement comme étant des paraspeckles (Cardinale et al 2007; Bond, & Fox 2009; Souquere et al 2010) . Les dimensions des paraspeckles observées par STORM dans les cellules GH4C1 sont inférieures à celles observées dans les paraspeckles de souris ou d'humain ce qui pourrait indiquer une variation de la taille des paraspeckles en fonction de l'espèce étudiée. Cependant sachant que le positionnement de Neat1, en particulier de la forme longue, semble déterminant pour la taille des paraspeckles (Souquere et al 2010), on peut postuler que dans notre modèle cellulaire le positionnement de Neat1-2 à l'intérieur des paraspeckles est différent de celui décrit dans d'autres types cellulaires. De plus, sachant que dans notre modèle cellulaire nous avons mis en évidence un rythme circadien du nombre de paraspeckles à l'intérieur du noyau, il reste à déterminer par analyse STORM si ce rythme s'accompagne ou non d'un rythme circadien de la taille et de la forme des paraspeckles.

## 3 - Neat1 et la rétention nucléaire des ARNs

#### Rétention d'ARNs par les paraspeckles dans la lignée GH4C1 :

Même si elles demeurent assez énigmatiques, une des fonctions les mieux caractérisées des paraspeckles concerne le rôle qu'ils jouent dans la rétention nucléaire des ARNs, empêchant leur export vers le cytoplasme et de ce fait leur traduction en protéines. A ce jour, les ARNs que l'on sait concernés par ce mécanisme de rétention nucléaire présentent dans leur région 3'UTR un motif IRAlu de la famille des séquences SINE reconnus par les paraspeckles (Prasanth et al 2005; Chen et al 2008; Mao et al 2011) . De fait en insérant ce motif IRAlu dans la région 3'UTR de l'ARNm de l'Egfp, l'ARN de l'egfp est plus fortement retenu dans le noyau et en corollaire l'expression de la protéine EGFP au niveau cytoplasmique est réduite.

En couplant la technique du RNA pull-down de Neat1 avec du séquençage haut débit des ARNs, nous avons cherché à identifier tous les ARNs endogènes qui sont cibles de Neat1. En utilisant deux sondes spécifiques vis-à-vis de Neat1 et en croisant les listes d'ARNs obtenues avec chacune de ces sondes, nous avons montré qu'un nombre élevé d'ARNs hypophysaires, environ 28% du transcriptome des GH4C1, s'associent à Neat1. Parmi eux, nous avons retrouvé les ARNm des deux hormones hypophysaires Gh et Prl exprimées dans les cellules GH4C1.

Le nombre important d'ARNs retrouvés associés à Neat1 suggère un rôle majeur de ce lncRNA dans la régulation du transcriptome hypophysaire avec potentiellement une implication forte dans les fonctions hormonales de l'hypophyse.

#### Etude des motifs permettant la rétention des ARNs par les paraspeckles

Différentes études ont permis de montrer que les paraspeckles reconnaissent et s'associent à des motifs de type IRAlu localisés dans la région 3'UTR des ARNs (Prasanth et al 2005; Chen et al 2008; Mao et al 2011) . Ces motifs IRAlu se composent d'éléments Alu de la famille des séquences SINE, répétés et inversés, qui s'apparient entre eux pour former des boucles d'ARN double brin. Ces boucles d'ARN double brin sont reconnues par des enzymes d'édition A -> I comme ADAR, et ce sont ces boucles

dans lesquelles les Adénosines sont éditées en Inosines qui seraient spécifiquement reconnues par un ou plusieurs éléments des paraspeckles. Chez les rongeurs, l'équivalent des motifs IRAlu correspond à des motifs IRSINE, composés de séquences SINE répétées et inversées formant comme dans les IRAlu des boucles double brin d'ARN.

A partir de la liste des ARNs trouvés associés aux paraspeckles dans la lignée GH4C1, nous avons cherché par analyse bioinformatique la présence de motifs de type IRSINE dans la région 3'UTR de ces ARNs. Etonnamment les régions 3'UTR des ARNs associés aux paraspeckles ne contiennent pas ce motif. Ces résultats suggèrent donc que d'autres motifs présents dans les ARNs permettent leur reconnaissance et leur association aux paraspeckles.

En sélectionnant trois ARNm dans la liste des ARNs associés aux paraspeckles, nous avons pu préciser que la région 3'UTR peut être en totalité ou en partie responsable de la liaison aux paraspeckles d'un ARNm. En effet, on a vu que pour l'ARNm de la Prl, des régions autres que la région 3'UTR, pourraient participer à cette association. En particulier, la région 5'UTR de l'ARNm de la Prl pourrait être impliquée dans cette association dans la mesure où il a été montré que deux des protéines du paraspeckles, NONO et SFPQ, peuvent en s'hétérodimérisant former un site de liaison ayant une forte affinité pour la région 5'UTR des ARNs (Duvignaud et al 2016). Il serait dès lors intéressant de déterminer si ce dimère protéique se lie à la région 5'UTR de l'ARNm de la Prl et si dans ce cas cette association implique les paraspeckles et entraîne une rétention nucléaire de l'ARNm.

Avec l'étude de l'implication de la région 3'UTR des ARNm de la Gh et de la Calr, nous avons de plus montré que plusieurs sous-domaines de cette région peuvent être collectivement impliqués dans l'association aux paraspeckles et nous suggérons que ces sous-domaines d'interaction pourraient correspondre aux différents motifs de liaison des nombreuses protéines dont sont composés les paraspeckles (Naganuma et al 2012).

Cette étude complémentaire nous permet donc de postuler d'une part que dans la région 3'UTR des ARNm il peut exister des motifs autres que les motifs IRAlu permettant l'association des ARNs aux paraspeckles et d'autre part, que des régions autres que la région 3'UTR en particulier la région 5'UTR pourraient être concernées par cette association.

Afin de poursuivre la caractérisation des éléments déterminants pour la reconnaissance des ARNs par les paraspeckles, la liste complète des ARNs associés aux paraspeckles sera analysée à l'aide d'outils bio-informatiques pour rechercher de manière globale l'existence de séquences répétées ou de motifs particuliers mais aussi de structures secondaires particulières que peuvent prendre ces ARNs. Un logiciel (RSAT-peak motifs) développé à Marseille par le Pr. Jacques Van Helden du laboratoire INSERM TAGC sera en particulier utilisé pour détecter des motifs communs aux différents ARNs. En complément, le logiciel BEAM permettra de rechercher des structures secondaires communes. Grâce à ces outils, il nous sera possible de rechercher à la fois dans les régions 3'UTR des ARNs mais également dans leurs régions 5'UTR la présence de motifs ou de structures secondaires communes aux ARNs identifiés comme des cibles des paraspeckles.

L'implication de ces motifs dans la liaison aux paraspeckles pourra alors être testée expérimentalement. Suivant la même approche que celle utilisée pour le motif IRAlu, les motifs identifiés par bio-informatique pourront être insérés dans la région 3'UTR de l'egfp et être exprimés dans les cellules GH4C1. Le RNA pull-down de Neat1 permettra alors de déterminer le rôle de ces motifs pour l'association des ARNs à Neat1. Les conséquences de l'insertion de ces motifs sur la rétention nucléaire de l'ARNm de l'egfp seront également évaluées.

## 4 – Les paraspeckles dans l'oscillateur hypophysaire

Rythmicité circadienne d'expression des éléments constituants des paraspeckles

Au cours d'études précédentes réalisées au sein de l'équipe, il a été montré que les protéines NONO et SFPQ, qui sont des constituants nécessaires à la formation des paraspeckles (Sasaki et al 2009), présentent une rythmicité circadienne d'expression dans les cellules hypophysaires GH4C1 (Guillaumond et al 2011; Becquet et al 2014). Notre étude complète ces données en montrant que cette rythmicité circadienne concerne également la protéine RBM14, qui est aussi un élément limitant dans la formation des paraspeckles ainsi que la protéine PSPC1 qui bien que ne jouant pas de rôle dans la formation des paraspeckles, est un des principaux constituants protéiques de ces structures. Il apparaît donc qu'au moins trois des protéines parmi les sept identifiées comme indispensables à la formation des paraspeckles (Naganuma et al 2012), présentent une rythmicité d'expression. De plus, l'élément structural majeur Neat1 et en particulier l'isoforme Neat1-2 dont dépend la formation des paraspeckles (Sasaki et al 2009), suivent également un rythme circadien d'expression. Cette rythmicité circadienne d'expression de Neat1 vient renforcer l'intérêt actuel de l'implication de lncRNAs dans le fonctionnement du système circadien. En effet, bon nombre de lncRNAs ont été impliqués dans le fonctionnement du système circadien via leur expression rythmique comme cela a été décrit dans la glande pinéale, un oscillateur endocrinien contrôlé par l'horloge centrale (Coon et al 2012). Dans d'autres oscillateurs circadiens comme le foie, il a été montré que l'expression rythmique de lncRNAs entraîne des modifications rythmiques de la chromatine (Vollmers et al 2012).

Une étude génomique réalisée récemment dans plusieurs organes de souris a montré qu'environ 32% des ARNs non codants les plus conservés sont exprimés selon un rythme circadien (Zhang et al 2014). Il est intéressant de souligner qu'au cours de notre étude nous avons pu montrer que l'expression de Neat1 est rythmique non seulement « in vitro » dans la lignée GH4C1 mais aussi « ex vivo » dans le tissu hypophysaire. Nous avons également observé un rythme d'expression de Neat1 dans l'horloge centrale, les noyaux suprachiasmatiques de l'hypothalamus, dans la rate et dans les glandes surrénales. Ces résultats pourraient suggérer que la rythmicité d'expression de Neat1 concerne tous les tissus de l'organisme. Cependant, les résultats de Zhang et al. semblent indiquer que les ARNs non codants rythment rarement dans tous les tissus de l'organisme, aucun d'entre eux n'oscillant dans plus de cinq organes différents (Zhang et al 2014). Il reste donc à identifier tous les tissus de l'organisme dans lesquels la formation des paraspeckles présente une rythmicité circadienne.

#### Conséquences de la rythmicité circadienne de formation des paraspeckles

La rythmicité du nombre de paraspeckles à l'intérieur du noyau des cellules peut avoir deux types de conséquences liées aux deux propriétés essentielles de ces structures qui sont d'une part la capacité à retenir dans le noyau les ARNs qui leur sont associés et d'autre part la capacité à séquestrer les protéines impliquées dans leur structure. Ainsi de par leur rythmicité les paraspeckles pourront contrôler la rythmicité des ARNs associés selon un mécanisme post-transcriptionnel et pourront également contrôler la rythmicité d'ARNs qui sont régulés au niveau transcriptionnel ou post-transcriptionnel par les différentes protéines qui entrent dans la composition de ces sous-domaines. En effet, différents auteurs ont montré que les paraspeckles en séquestrant les protéines qui leur sont associées peuvent contrôler l'activité de ces protéines (Imamura et al 2014; Hirose et al 2013) . Dans leur étude, Naganuma et al. montrent que les paraspeckles peuvent ainsi séquestrer une trentaine de protéines qui sont majoritairement impliquées dans des mécanismes transcriptionnels et posttranscriptionnels de régulation de l'expression des gènes (Naganuma et al 2012). La rythmicité d'expression des paraspeckles peut donc conduire à une séquestration rythmique des protéines par les paraspeckles, et dès lors à une rythmicité de leur activité.

La rythmicité du nombre de paraspeckles permet aussi de contrôler la rythmicité des ARNs qui leur sont associés. De fait, nous avons pu montrer que la construction IRAluegfp ainsi que plusieurs ARNs endogènes ont une rétention nucléaire rythmique qui est dépendante des paraspeckles. Notre analyse des ARNs hypophysaires associés à Neat1 ayant montré l'importance quantitative de ces ARNs, la rythmicité de formation des paraspeckles pourrait contribuer de façon notable à la rythmicité circadienne du transcriptome hypophysaire.

Au final, il apparaît que la rythmicité de formation des paraspeckles que nous avons décrite possède le potentiel de réguler de manière directe ou indirecte via les protéines de ces structures un nombre important d'ARNs ce qui pourrait témoigner d'un rôle majeur des paraspeckles dans la régulation des fonctions circadiennes.

Dans le but de préciser le nombre d'ARNs dont la rythmicité d'expression dépend des paraspeckles, nous avons développé plusieurs lignées clonales de cellules GH4C1 KO Neat1, dépourvues de paraspeckles, que nous avons établies en utilisant la technique CRISPR/Cas9 (Ran et al 2013). La méthode Crispr-Cas9 utilise des guides oligonucléotidiques qui ciblent la région de l'ADN à éditer. Ces sondes vont permettre la fixation sur cette région de la protéine nucléase Cas9 responsable du clivage de l'ADN. Par ce mécanisme, une séquence correspondant à Neat1-1 ainsi qu'à la partie 5' de Neat1-2 a été supprimée, induisant la perte totale d'expression des deux isoformes. L'analyse comparative des ARNs rythmiques dans les lignées KO par rapport à la lignée contrôle permettra de déterminer quelle proportion du transcriptome rythmique hypophysaire dépend de Neat1.

Des travaux récents ont montré qu'outre son implication dans la formation et les fonctions des paraspeckles, Neat1 pouvait aussi intervenir dans le contrôle de la transcription des gènes en se liant à la chromatine à proximité des sites de début ou de fin de transcription (West et al 2014). La régulation de la rythmicité des gènes par Neat1 pourrait donc se faire également au niveau transcriptionnel. Par conséquent, nous rechercherons dans les cellules GH4C1 les sites de liaison de Neat1 à l'ADN en utilisant la technique CHART (capture hybridization analysis of RNA targets (Simon 2013)). Le CHART, tout comme le RNA pull-down, utilise des sondes oligonucléotidiques spécifiques de Neat1. A la différence du RNA pull-down, le CHART va permettre d'entrainer non pas les ARNs associés à Neat1 mais les différentes régions de la chromatine où il est fixé (Simon 2013). En couplant les expériences de CHART à du séquençage haut débit d'ADN, nous pourrons évaluer les gènes qui peuvent être des cibles transcriptionnelles potentielles de Neat1. Cette liste de gènes sera confrontée à celle des gènes dont l'expression est rythmique dans les cellules GH4C1.

Finalement, nous pourrons comparer la liste des ARNs dont la rythmicité dépend de Neat1 que nous aurons établie à partir de nos lignées clonales KO Neat1, avec la liste des gènes qui sont des cibles transcriptionnelles potentielles de Neat1 ainsi qu'avec la liste des gènes dont les ARNs sont associés aux paraspeckles. Dès lors, l'ensemble de cette étude nous permettra de préciser la contribution de Neat1 au transcriptome rythmique des cellules hypophysaires via un contrôle post-transcriptionnel impliquant les paraspeckles et la rétention nucléaire des ARNm et/ou via un contrôle transcriptionnel.

## 5 – Neat1, organogénèse et rythmicité circadienne :

Un modèle murin KO pour Neat1, a permis de montrer l'implication de Neat1 dans le contrôle du développement de certains organes. Deux études réalisées sur ce modèle murin montrent en effet que Neat1 est impliqué d'une part dans le développement d'un corps jaune fonctionnel (Nakagawa et al 2014) et d'autre part dans le développement de la glande mammaire chez les femelles allaitantes (Standaert et al 2014). Les mécanismes par lesquels Neat1 participe au développement du corps jaune sont encore inconnus. Au niveau du développement de la glande mammaire pendant la lactation, l'absence d'expression de Neat1 induit une réduction de la prolifération des cellules épithéliales mammaires qui se traduit par un sous développement de la glande et par conséquent une réduction de la quantité de lait produite lors de l'allaitement.

Cette implication de Neat1 dans la prolifération cellulaire et l'organogénèse est cohérente avec l'observation d'une dérégulation de son expression dans de nombreux cancers (Gibb et al 2011). Dans les gliomes, le cancer du sein, des poumons, de l'œsophage, de l'estomac, du foie, des ovaires et de la prostate, l'expression de Neat1 augmente en corrélation avec une augmentation de la taille des tumeurs et de leurs métastases (Lo et al 2016).

On sait qu'il existe des liens forts entre l'horloge circadienne et les processus de prolifération et de croissance cellulaires. L'horloge circadienne contrôle notamment l'expression de nombreux acteurs du cycle cellulaire et des perturbations de l'oscillateur circadien induisent des dérégulations de la division cellulaire et une augmentation de la tumorigénèse (Shostak 2017; Kiessling et al 2017). Il est dès lors tentant de postuler que la rythmicité d'expression de Neat1 que nous rapportons peut être le support de son implication dans les processus de prolifération et d'organogénèse. Un autre long ARN non-codant lui aussi impliqué dans le fonctionnement du système circadien illustre la forte corrélation qui existe entre les processus de prolifération cellulaire, en particulier au cours de différents types de cancer et la dérégulation de lncRNAs. En effet, le lncRNA Highly Upregulated in Liver Cancer (HULC), dont l'expression est augmentée dans les carcinomes hépatiques, accélère la carcinogénèse en altérant l'oscillateur circadien hépatique par stimulation de l'expression de la protéine CLOCK (Cui et al 2015).

On sait que l'horloge circadienne joue un rôle majeur dans le développement de la glande mammaire. En effet des modèles murins de perturbation des deux principaux l'oscillateur moléculaire. Bmal1 acteurs de et Clock. montrent qu'un dysfonctionnement du système circadien est corrélé à un défaut de développement de la glande mammaire (Dolatshad et al 2006; Metz et al 2006; Boden et al 2010; Casey et al 2016) . Ces modèles murins à l'instar du modèle KO Neat1, présentent une réduction de la longueur et de l'arborisation des canaux à l'intérieur de la glande, un déficit de croissance qui se traduit au final par une perturbation de la production de lait. Il est alors tentant de postuler que la rythmicité de Neat1 peut jouer un rôle dans la régulation de la formation de la glande mammaire. Il est cependant impossible de déterminer à partir des études précédemment citées si les défauts de développement de la glande mammaire rapportés résultent de la perturbation locale de l'oscillateur moléculaire de la glande mammaire, ou si ces défauts sont la conséquence de la perturbation d'oscillateurs dans d'autres tissus. Nous savons par notre étude que dans les cellules hypophysaires, les ARNm de la Gh et de la Prl sont associés aux paraspeckles et qu'ils sont donc potentiellement dérégulés en cas de KO Neat1. Par conséquent, ces deux hormones jouant un rôle majeur dans le développement de la glande mammaire (NANDI 1958; LYONS 1958; Haslam 1989; Brisken et al 1999; Gallego et al 2001; Mallepell et al 2006) , il est envisageable que les perturbations du développement de la glande mammaire dans le modèle souris KO Neat1 proviennent au moins partiellement de la perturbation de l'expression de ces hormones liée à la perte de Neat1 dans l'hypophyse.

La description de la rythmicité d'expression de Neat1 nous permet donc de postuler l'existence d'autres mécanismes par lesquels Neat1 pourrait contrôler la prolifération cellulaire et l'organogénèse.

## **Conclusion générale**

Au cours de cette étude nous avons montré que Neat1 et les paraspeckles présentent « in vitro » dans une lignée de cellules hypophysaires mais également « ex vivo » dans différents oscillateurs dont l'horloge centrale, une rythmicité circadienne d'expression. Cette rythmicité du nombre de paraspeckles à l'intérieur du noyau des cellules va permettre le contrôle post-transcriptionnel rythmique de nombreux ARNs rythmiquement retenus dans le noyau. La mise au point d'une technique de Neat1 RNA pull-down suivie d'un séquençage haut débit des ARNs nous a permis d'identifier un grand nombre de ces ARNs associés aux paraspeckles, suggérant que par ce mécanisme de contrôle post-transcriptionnel, Neat1 pourrait être un régulateur majeur de l'expression circadienne des gènes dans les structures où son expression est rythmique.

Cependant la rythmicité circadienne de Neat1 que nous avons décrite pourrait avoir d'autres conséquences fonctionnelles liées à la diversité des mécanismes par lesquels Neat1 exerce ses effets. En particulier, on sait que Neat1 via les paraspeckles, est impliqué non seulement dans la rétention des ARNs qui leur sont associés mais aussi dans la séquestration des protéines architecturales de ces sous-domaines (Naganuma et al 2012; Hirose et al 2013). De ce fait la séquestration rythmique de ces protéines affectera la rythmicité de toutes les voies de régulation dans lesquelles ces protéines sont engagées. Finalement, parce que Neat1 possède de nombreux sites de liaison à la chromatine (West et al 2014), sa liaison rythmique à l'ADN pourrait être engagée dans une régulation transcriptionnelle rythmique de gènes.

Au final, la rythmicité circadienne de Neat1 conjuguée à la diversité de ses mécanismes d'action pourraient faire de ce lncRNA un acteur majeur du fonctionnement du système circadien.

Sachant que l'étude de modèles murins KO Neat1 a permis d'impliquer ce lncRNA dans les processus de prolifération cellulaire et d'organogénèse (Standaert et al 2014; Nakagawa et al 2014), des processus en étroite association avec le fonctionnement du système circadien, notre étude soulève la question du rôle possible de la rythmicité de Neat1 dans ces processus. Dès lors, en apportant de nouvelles données sur la contribution de Neat1 au fonctionnement du système circadien, notre travail offre de nouveaux éléments de compréhension de la façon dont Neat1 pourrait être impliqué non seulement au plan physiologique dans l'organogénèse, mais aussi au plan pathologique dans la tumorigenèse.

## **Bibliographie**

- Abe, M., Herzog, E.D., Yamazaki, S., Straume, M., Tei, H., Sakaki, Y., Menaker, M. & Block,
  G.D., 2002, Circadian rhythms in isolated brain regions, *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 22(1), pp. 350-6.
- Abraham, U., Granada, A.E., Westermark, P.O., Heine, M., Kramer, A. & Herzel, H., 2010, Coupling governs entrainment range of circadian clocks, *Mol Syst Biol*, 6, p. 438.
- Abrahamson, E.E. & Moore, R.Y., 2001, Suprachiasmatic nucleus in the mouse: retinal innervation, intrinsic organization and efferent projections, *Brain research*, 916(1), pp. 172-91.
- Aioun, J., Chambille, I., Peytevin, J. & Martinet, L., 1998, Neurons containing gastrinreleasing peptide and vasoactive intestinal polypeptide are involved in the reception of the photic signal in the suprachiasmatic nucleus of the Syrian hamster: an immunocytochemical ultrastructural study, *Cell and tissue research*, 291(2), pp. 239-53.
- Akashi, M. & Nishida, E., 2000, Involvement of the MAP kinase cascade in resetting of the mammalian circadian clock, *Genes & development*, 14(6), pp. 645-9.
- Akashi, M. & Takumi, T., 2005, The orphan nuclear receptor RORalpha regulates circadian transcription of the mammalian core-clock Bmal1, *Nature structural & molecular biology*, 12(5), pp. 441-8.
- Akashi, M., Tsuchiya, Y., Yoshino, T. & Nishida, E., 2002, Control of intracellular dynamics of mammalian period proteins by casein kinase I varepsilon (CKIvarepsilon) and CKIδ in cultured cells, *Molecular and cellular biology*, 22(6), pp. 1693-703.
- Akhtar, R.A., Reddy, A.B., Maywood, E.S., Clayton, J.D., King, V.M., Smith, A.G., Gant, T.W., Hastings, M.H. & Kyriacou, C.P., 2002, Circadian cycling of the mouse liver transcriptome, as revealed by cDNA microarray, is driven by the suprachiasmatic nucleus, *Current biology : CB*, 12(7), pp. 540-50.
- Akiyama, M., Kouzu, Y., Takahashi, S., Wakamatsu, H., Moriya, T., Maetani, M., Watanabe, S., Tei, H., Sakaki, Y. & Shibata, S., 1999, Inhibition of light- or glutamate-induced mPer1 expression represses the phase shifts into the mouse circadian locomotor and suprachiasmatic firing rhythms, *The Journal of*

*neuroscience : the official journal of the Society for Neuroscience*, 19(3), pp. 1115-21.

- Albers, H.E. & Ferris, C.F., 1984, Neuropeptide Y: role in light-dark cycle entrainment of hamster circadian rhythms, *Neuroscience letters*, 50(1-3), pp. 163-8.
- Alvarez, J.D. & Sehgal, A., 2005, The thymus is similar to the testis in its pattern of circadian clock gene expression, *Journal of biological rhythms*, 20(2), pp. 111-21.
- Anantharaman, A., Jadaliha, M., Tripathi, V., Nakagawa, S., Hirose, T., Jantsch, M.F., Prasanth, S.G. & Prasanth, K.V., 2016, Paraspeckles modulate the intranuclear distribution of paraspeckle-associated Ctn RNA, *Scientific reports*, 6, p. 34043.
- Atger, F., Gobet, C., Marquis, J., Martin, E., Wang, J., Weger, B., Lefebvre, G., Descombes,
  P., Naef, F. & Gachon, F., 2015, Circadian and feeding rhythms differentially affect rhythmic mRNA transcription and translation in mouse liver, *Proceedings of the National Academy of Sciences of the United States of America*.
- Balsalobre, A., Damiola, F. & Schibler, U., 1998, A serum shock induces circadian gene expression in mammalian tissue culture cells, *Cell*, 93(6), pp. 929-37.
- Becquet, D., Boyer, B., Rasolonjanahary, R., Brue, T., Guillen, S., Moreno, M., Franc, J.L. & François-Bellan, A.M., 2014, Evidence for an internal and functional circadian clock in rat pituitary cells, *Mol Cell Endocrinol*, 382(2), pp. 888-98.
- Bell-Pedersen, D., Cassone, V.M., Earnest, D.J., Golden, S.S., Hardin, P.E., Thomas, T.L. & Zoran, M.J., 2005, Circadian rhythms from multiple oscillators: lessons from diverse organisms, *Nat Rev Genet*, 6(7), pp. 544-56.
- Berk, M.L. & Finkelstein, J.A., 1981, An autoradiographic determination of the efferent projections of the suprachiasmatic nucleus of the hypothalamus, *Brain research*, 226(1), pp. 1-13.
- Biamonti, G. & Vourc'h, C., 2010, Nuclear stress bodies, *Cold Spring Harbor perspectives in biology*, 2(6), p. a000695.
- Boden, M.J., Varcoe, T.J., Voultsios, A. & Kennaway, D.J., 2010, Reproductive biology of female Bmal1 null mice, *Reproduction (Cambridge, England)*, 139(6), pp. 1077-90.
- Bond, C.S. & Fox, A.H., 2009, Paraspeckles: nuclear bodies built on long noncoding RNA, *The Journal of cell biology*, 186(5), p. 637.

- Bosler, O., 1989, Ultrastructural relationships of serotonin and GABA terminals in the rat suprachiasmatic nucleus. Evidence for a close interconnection between the two afferent systems, *Journal of neurocytology*, 18(1), pp. 105-13.
- Bosler, O. & Beaudet, A., 1985, VIP neurons as prime synaptic targets for serotonin afferents in rat suprachiasmatic nucleus: a combined radioautographic and immunocytochemical study, *Journal of neurocytology*, 14(5), pp. 749-63.
- Brisken, C., Kaur, S., Chavarria, T.E., Binart, N., Sutherland, R.L., Weinberg, R.A., Kelly,
  P.A. & Ormandy, C.J., 1999, Prolactin controls mammary gland development via direct and indirect mechanisms, *Developmental biology*, 210(1), pp. 96-106.
- Brown, C.J., Hendrich, B.D., Rupert, J.L., Lafrenière, R.G., Xing, Y., Lawrence, J. & Willard, H.F., 1992, The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus, *Cell*, 71(3), pp. 527-42.
- Brown, C.J., Lafreniere, R.G., Powers, V.E., Sebastio, G., Ballabio, A., Pettigrew, A.L., Ledbetter, D.H., Levy, E., Craig, I.W. & Willard, H.F., 1991, Localization of the X inactivation centre on the human X chromosome in Xq13, *Nature*, 349(6304), pp. 82-4.
- Brown, J.A., Valenstein, M.L., Yario, T.A., Tycowski, K.T. & Steitz, J.A., 2012, Formation of triple-helical structures by the 3'-end sequences of MALAT1 and MENβ noncoding RNAs, *Proceedings of the National Academy of Sciences of the United States of America*, 109(47), pp. 19202-7.
- Buhr, E.D., Yoo, S.H. & Takahashi, J.S., 2010, Temperature as a universal resetting cue for mammalian circadian oscillators, *Science (New York, N.Y.)*, 330(6002), pp. 379-85.
- Buijs, R.M. & Kalsbeek, A., 2001, Hypothalamic integration of central and peripheral clocks, *Nature reviews. Neuroscience*, 2(7), pp. 521-6.
- Bunger, M.K., Wilsbacher, L.D., Moran, S.M., Clendenin, C., Radcliffe, L.A., Hogenesch, J.B., Simon, M.C., Takahashi, J.S. & Bradfield, C.A., 2000, Mop3 is an essential component of the master circadian pacemaker in mammals, *Cell*, 103(7), pp. 1009-17.
- Capshew, C.R., Dusenbury, K.L. & Hundley, H.A., 2012, Inverted Alu dsRNA structures do not affect localization but can alter translation efficiency of human mRNAs independent of RNA editing, *Nucleic acids research*, 40(17), pp. 8637-45.

- Cardinale, S., Cisterna, B., Bonetti, P., Aringhieri, C., Biggiogera, M. & Barabino, S.M., 2007, Subnuclear localization and dynamics of the Pre-mRNA 3' end processing factor mammalian cleavage factor I 68-kDa subunit, *Molecular biology of the cell*, 18(4), pp. 1282-92.
- Casey, T., Crodian, J., Suárez-Trujillo, A., Erickson, E., Weldon, B., Crow, K., Cummings, S., Chen, Y., Shamay, A., Mabjeesh, S.J. & Plaut, K., 2016, CLOCK Regulates Mammary Epithelial Cell Growth and Differentiation, *American journal of physiology*. *Regulatory, integrative and comparative physiology*, pp. ajpregu.00032.2016.
- Cermakian, N. & Sassone-Corsi, P., 2002, Environmental stimulus perception and control of circadian clocks, *Current opinion in neurobiology*, 12(4), pp. 359-65.
- Cermakian, N., Monaco, L., Pando, M.P., Dierich, A. & Sassone-Corsi, P., 2001, Altered behavioral rhythms and clock gene expression in mice with a targeted mutation in the Period1 gene, *EMBO J*, 20(15), pp. 3967-74.
- Chen, L.L. & Carmichael, G.G., 2008, Gene regulation by SINES and inosines: biological consequences of A-to-I editing of Alu element inverted repeats, *Cell cycle (Georgetown, Tex.)*, 7(21), pp. 3294-301.
- Chen, L.L. & Carmichael, G.G., 2009, Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA, *Molecular cell*, 35(4), pp. 467-78.
- Chen, L.L., DeCerbo, J.N. & Carmichael, G.G., 2008, Alu element-mediated gene silencing, *EMBO J*, 27(12), pp. 1694-705.
- Christian, C.A., Mobley, J.L. & Moenter, S.M., 2005, Diurnal and estradiol-dependent changes in gonadotropin-releasing hormone neuron firing activity, *Proceedings of the National Academy of Sciences of the United States of America*, 102(43), pp. 15682-7.
- Chu, A., Zhu, L., Blum, I.D., Mai, O., Leliavski, A., Fahrenkrug, J., Oster, H., Boehm, U. & Storch, K.F., 2013, Global but not gonadotrope-specific disruption of Bmal1 abolishes the luteinizing hormone surge without affecting ovulation, *Endocrinology*, 154(8), pp. 2924-35.
- Chu, C., Qu, K., Zhong, F.L., Artandi, S.E. & Chang, H.Y., 2011, Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions, *Molecular cell*, 44(4), pp. 667-78.

- Chu, C., Quinn, J. & Chang, H.Y., 2012, Chromatin isolation by RNA purification (ChIRP), *J Vis Exp*, 61, p. e3912.
- Clemson, C.M., Hutchinson, J.N., Sara, S.A., Ensminger, A.W., Fox, A.H., Chess, A. & Lawrence, J.B., 2009, An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles, *Molecular cell*, 33(6), pp. 717-26.
- Coon, S.L., Munson, P.J., Cherukuri, P.F., Sugden, D., Rath, M.F., Møller, M., Clokie, S.J., Fu,
  C., Olanich, M.E., Rangel, Z., Werner, T., Mullikin, J.C., Klein, D.C. & NISC
  Comparative Sequencing Program, 2012, Circadian changes in long noncoding
  RNAs in the pineal gland, *Proceedings of the National Academy of Sciences of the*United States of America, 109(33), pp. 13319-24.
- Critchlow, V., 1963, The role of light in the neuroendocrine system, *Advances in neuroendocrinology*, pp. 377-402.
- Cui, M., Zheng, M., Sun, B., Wang, Y., Ye, L. & Zhang, X., 2015, A long noncoding RNA perturbs the circadian rhythm of hepatoma cells to facilitate hepatocarcinogenesis, *Neoplasia*, 17(1), pp. 79-88.
- Czeisler, C.A. & Klerman, E.B., 1999, Circadian and sleep-dependent regulation of hormone release in humans, *Recent progress in hormone research*, 54, pp. 97-130; discussion 130-2.
- Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., Guernec, G., Martin, D., Merkel, A., Knowles, D.G., Lagarde, J., Veeravalli, L., Ruan, X., Ruan, Y., Lassmann, T., Carninci, P., Brown, J.B., Lipovich, L., Gonzalez, J.M., Thomas, M., Davis, C.A., Shiekhattar, R., Gingeras, T.R., Hubbard, T.J., Notredame, C., Harrow, J. & Guigó, R., 2012, The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression, *Genome research*, 22(9), pp. 1775-89.
- Dibner, C., Sage, D., Unser, M., Bauer, C., d'Eysmond, T., Naef, F. & Schibler, U., 2009, Circadian gene expression is resilient to large fluctuations in overall transcription rates, *EMBO J*, 28(2), pp. 123-34.
- Dibner, C., Schibler, U. & Albrecht, U., 2010, The mammalian circadian timing system: organization and coordination of central and peripheral clocks, *Annu Rev Physiol*, 72, pp. 517-49.

- Dimitrova, N., Zamudio, J.R., Jong, R.M., Soukup, D., Resnick, R., Sarma, K., Ward, A.J., Raj, A., Lee, J.T., Sharp, P.A. & Jacks, T., 2014, LincRNA-p21 activates p21 in cis to promote Polycomb target gene expression and to enforce the G1/S checkpoint, *Molecular cell*, 54(5), pp. 777-90.
- Ding, J.M., Faiman, L.E., Hurst, W.J., Kuriashkina, L.R. & Gillette, M.U., 1997, Resetting the biological clock: mediation of nocturnal CREB phosphorylation via light, glutamate, and nitric oxide, *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 17(2), pp. 667-75.
- Dolatshad, H., Campbell, E.A., O'Hara, L., Maywood, E.S., Hastings, M.H. & Johnson, M.H., 2006, Developmental and reproductive performance in circadian mutant mice, *Human reproduction (Oxford, England)*, 21(1), pp. 68-79.
- Drucker-Colín, R., Aguilar-Roblero, R., García-Hernández, F., Fernández-Cancino, F. & Bermudez Rattoni, F., 1984, Fetal suprachiasmatic nucleus transplants: diurnal rhythm recovery of lesioned rats, *Brain research*, 311(2), pp. 353-7.
- Duffield, G.E., 2003, DNA microarray analyses of circadian timing: the genomic basis of biological time, *Journal of neuroendocrinology*, 15(10), pp. 991-1002.
- Dunn, J.D., Arimura, A. & Scheving, L.E., 1972, Effect of stress on circadian periodicity in serum LH and prolactin concentration, *Endocrinology*, 90(1), pp. 29-33.
- Duvignaud, J.B., Bédard, M., Nagata, T., Muto, Y., Yokoyama, S., Gagné, S.M. & Vincent, M.,
   2016, Structure, dynamics and interaction of p54nrb/NonO RRM1 with 5' Splice
   Site RNA sequence, *Biochemistry*.
- Edgar, D.M., Miller, J.D., Prosser, R.A., Dean, R.R. & Dement, W.C., 1993, Serotonin and the mammalian circadian system: II. Phase-shifting rat behavioral rhythms with serotonergic agonists, *Journal of biological rhythms*, 8(1), pp. 17-31.
- Eide, E.J., Vielhaber, E.L., Hinz, W.A. & Virshup, D.M., 2002, The circadian regulatory proteins BMAL1 and cryptochromes are substrates of casein kinase lepsilon, *The Journal of biological chemistry*, 277(19), pp. 17248-54.
- Eide, E.J., Woolf, M.F., Kang, H., Woolf, P., Hurst, W., Camacho, F., Vielhaber, E.L., Giovanni, A. & Virshup, D.M., 2005, Control of mammalian circadian rhythm by CKIepsilon-regulated proteasome-mediated PER2 degradation, *Molecular and cellular biology*, 25(7), pp. 2795-807.
- Elbarbary, R.A. & Maquat, L.E., 2015, CARMing down the SINEs of anarchy: two paths to freedom from paraspeckle detention, *Genes & development*, 29(7), pp. 687-9.

- ENCODE Project Consortium, 2012, An integrated encyclopedia of DNA elements in the human genome, *Nature*, 489(7414), pp. 57-74.
- Engreitz, J.M., Sirokman, K., McDonel, P., Shishkin, A.A., Surka, C., Russell, P., Grossman, S.R., Chow, A.Y., Guttman, M. & Lander, E.S., 2014, RNA-RNA interactions enable specific targeting of noncoding RNAs to nascent Pre-mRNAs and chromatin sites, *Cell*, 159(1), pp. 188-99.
- Evans, J.A., Pan, H., Liu, A.C. & Welsh, D.K., 2012, Cry1-/- circadian rhythmicity depends on SCN intercellular coupling, *Journal of biological rhythms*, 27(6), pp. 443-52.
- Faghihi, M.A., Modarresi, F., Khalil, A.M., Wood, D.E., Sahagan, B.G., Morgan, T.E., Finch, C.E., St Laurent, G., Kenny, P.J. & Wahlestedt, C., 2008, Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of beta-secretase, *Nat Med*, 14(7), pp. 723-30.
- Farnell, Y.F., Shende, V.R., Neuendorff, N., Allen, G.C. & Earnest, D.J., 2011, Immortalized cell lines for real-time analysis of circadian pacemaker and peripheral oscillator properties, *The European journal of neuroscience*, 33(8), pp. 1533-40.
- Feldman, J.F., 1967, Lengthening the period of a biological clock in Euglena by cycloheximide, an inhibitor of protein synthesis, *Proceedings of the National Academy of Sciences*, 57(4), pp. 1080-7.
- Fitzpatrick, T. & Huang, S., 2012, 3'-UTR-located inverted Alu repeats facilitate mRNA translational repression and stress granule accumulation, *Nucleus (Austin, Tex.)*, 3(4), pp. 359-69.
- Fox, A.H., Bond, C.S. & Lamond, A.I., 2005, P54nrb forms a heterodimer with PSP1 that localizes to paraspeckles in an RNA-dependent manner, *Molecular biology of the cell*, 16(11), pp. 5304-15.
- Fox, A.H., Lam, Y.W., Leung, A.K., Lyon, C.E., Andersen, J., Mann, M. & Lamond, A.I., 2002, Paraspeckles: a novel nuclear domain, *Current biology : CB*, 12(1), pp. 13-25.
- Francois-Bellan, A.M., Kachidian, P., Dusticier, G., Tonon, M.C., Vaudry, H. & Bosler, O., 1990, GABA neurons in the rat suprachiasmatic nucleus: involvement in chemospecific synaptic circuitry and evidence for GAD-peptide colocalization, *Journal of neurocytology*, 19(6), pp. 937-47.
- Fukuhara, C., Brewer, J.M., Dirden, J.C., Bittman, E.L., Tosini, G. & Harrington, M.E., 2001, Neuropeptide Y rapidly reduces Period 1 and Period 2 mRNA levels in the hamster suprachiasmatic nucleus, *Neuroscience letters*, 314(3), pp. 119-22.

- Gallego, M.I., Binart, N., Robinson, G.W., Okagaki, R., Coschigano, K.T., Perry, J., Kopchick,
  J.J., Oka, T., Kelly, P.A. & Hennighausen, L., 2001, Prolactin, growth hormone, and
  epidermal growth factor activate Stat5 in different compartments of mammary
  tissue and exert different and overlapping developmental effects, *Developmental biology*, 229(1), pp. 163-75.
- Gau, D., Lemberger, T., von Gall, C., Kretz, O., Le Minh, N., Gass, P., Schmid, W., Schibler,
  U., Korf, H.W. & Schütz, G., 2002, Phosphorylation of CREB Ser142 regulates
  light-induced phase shifts of the circadian clock, *Neuron*, 34(2), pp. 245-53.
- Gekakis, N., Staknis, D., Nguyen, H.B., Davis, F.C., Wilsbacher, L.D., King, D.P., Takahashi, J.S. & Weitz, C.J., 1998, Role of the CLOCK protein in the mammalian circadian mechanism, *Science*, 280(5369), pp. 1564-9.
- Gibb, E.A., Vucic, E.A., Enfield, K.S., Stewart, G.L., Lonergan, K.M., Kennett, J.Y., Becker-Santos, D.D., MacAulay, C.E., Lam, S., Brown, C.J. & Lam, W.L., 2011, Human cancer long non-coding RNA transcriptomes, *PloS one*, 6(10), p. e25915.
- Ginty, D.D., Kornhauser, J.M., Thompson, M.A., Bading, H., Mayo, K.E., Takahashi, J.S. & Greenberg, M.E., 1993, Regulation of CREB phosphorylation in the suprachiasmatic nucleus by light and a circadian clock, *Science*, 260(5105), pp. 238-41.
- Girotti, M., Weinberg, M.S. & Spencer, R.L., 2009, Diurnal expression of functional and clock-related genes throughout the rat HPA axis: system-wide shifts in response to a restricted feeding schedule, *American journal of physiology. Endocrinology and metabolism*, 296(4), pp. E888-97.
- Glossop, N.R.J., Lyons, L.C. & Hardin, P.E., 1999, Interlocked Feedback Loops Within the Drosophila Circadian Oscillator, *Science (New York, N.Y.)*, 286(5440), pp. 766-8.
- Golombek, D.A., Biello, S.M., Rendon, R.A. & Harrington, M.E., 1996, Neuropeptide Y phase shifts the circadian clock in vitro via a Y2 receptor, *Neuroreport*, 7(7), pp. 1315-9.
- Grattan, D.R. & Kokay, I.C., 2008, Prolactin: a pleiotropic neuroendocrine hormone, *Journal of neuroendocrinology*, 20(6), pp. 752-63.
- Green, C.B., 2017, Circadian Posttranscriptional Regulatory Mechanisms in Mammals, *Cold Spring Harbor perspectives in biology*.
- Green, D.J. & Gillette, R., 1982, Circadian rhythm of firing rate recorded from single cells in the rat suprachiasmatic brain slice, *Brain research*, 245(1), pp. 198-200.

- Guillaumond, F., Boyer, B., Becquet, D., Guillen, S., Kuhn, L., Garin, J., Belghazi, M., Bosler,
  O., Franc, J.L. & François-Bellan, A.M., 2011, Chromatin remodeling as a mechanism for circadian prolactin transcription: rhythmic NONO and SFPQ recruitment to HLTF, *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 25(8), pp. 2740-56.
- Guillaumond, F., Dardente, H., Giguère, V. & Cermakian, N., 2005, Differential control of Bmal1 circadian transcription by REV-ERB and ROR nuclear receptors, *Journal of biological rhythms*, 20(5), pp. 391-403.
- Guo, H., Guo, H., Brewer, J.M., Lehman, M.N. & Bittman, E.L., 2006, Suprachiasmatic regulation of circadian rhythms of gene expression in hamster peripheral organs: effects of transplanting the pacemaker, *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 26(24), pp. 6406-12.
- Harmer, S.L., Panda, S. & Kay, S.A., 2001, Molecular bases of circadian rhythms, *Annual review of cell and developmental biology*, 17(1), pp. 215-53.
- Harms, E., Kivimäe, S., Young, M.W. & Saez, L., 2004, Posttranscriptional and posttranslational regulation of clock genes, *Journal of biological rhythms*, 19(5), pp. 361-73.
- Harrington, M.E. & Schak, K.M., 2000, Neuropeptide Y phase advances the in vitro hamster circadian clock during the subjective day with no effect on phase during the subjective night, *Canadian journal of physiology and pharmacology*, 78(2), pp. 87-92.
- Haslam, S.Z., 1989, The ontogeny of mouse mammary gland responsiveness to ovarian steroid hormones, *Endocrinology*, 125(5), pp. 2766-72.
- Herzog, E.D., Aton, S.J., Numano, R., Sakaki, Y. & Tei, H., 2004, Temporal precision in the mammalian circadian system: a reliable clock from less reliable neurons, *Journal of biological rhythms*, 19(1), pp. 35-46.
- Herzog, E.D., Takahashi, J.S. & Block, G.D., 1998, Clock controls circadian period in isolated suprachiasmatic nucleus neurons, *Nat Neurosci*, 1(8), pp. 708-13.
- Hirose, T., Virnicchi, G., Tanigawa, A., Naganuma, T., Li, R., Kimura, H., Yokoi, T., Nakagawa, S., Bénard, M., Fox, A.H. & Pierron, G., 2013, NEAT1 long noncoding RNA regulates transcription via protein sequestration within subnuclear bodies, *Molecular biology of the cell*, 25(1), pp. 169-83.
- Hisano, S., Chikamori-Aoyama, M., Katoh, S., Kagotani, Y., Daikoku, S. & Chihara, K., 1988, Suprachiasmatic nucleus neurons immunoreactive for vasoactive intestinal polypeptide have synaptic contacts with axons immunoreactive for neuropeptide Y: an immunoelectron microscopic study in the rat, *Neuroscience letters*, 88(2), pp. 145-50.
- Hood, S. & Amir, S., 2017, Neurodegeneration and the Circadian Clock, *Frontiers in aging neuroscience*, 9, p. 170.
- Hosoda, H., Motohashi, J., Kato, H., Masushige, S. & Kida, S., 2004, A BMAL1 mutant with arginine 91 substituted with alanine acts as a dominant negative inhibitor, *Gene*, 338(2), pp. 235-41.
- Hu, S.B., Xiang, J.F., Li, X., Xu, Y., Xue, W., Huang, M., Wong, C.C., Sagum, C.A., Bedford, M.T., Yang, L., Cheng, D. & Chen, L.L., 2015, Protein arginine methyltransferase CARM1 attenuates the paraspeckle-mediated nuclear retention of mRNAs containing IRAlus, *Genes & development*, 29(6), pp. 630-45.
- Huarte, M., Guttman, M., Feldser, D., Garber, M., Koziol, M.J., Kenzelmann-Broz, D.,
  Khalil, A.M., Zuk, O., Amit, I., Rabani, M., Attardi, L.D., Regev, A., Lander, E.S.,
  Jacks, T. & Rinn, J.L., 2010, A large intergenic noncoding RNA induced by p53
  mediates global gene repression in the p53 response, *Cell*, 142(3), pp. 409-19.
- Hughes, M., Deharo, L., Pulivarthy, S.R., Gu, J., Hayes, K., Panda, S. & Hogenesch, J.B.,
  2007, High-resolution time course analysis of gene expression from pituitary, *Cold Spring Harbor symposia on quantitative biology*, 72, pp. 381-6.
- Hughes, M.E., DiTacchio, L., Hayes, K.R., Vollmers, C., Pulivarthy, S., Baggs, J.E., Panda, S.
  & Hogenesch, J.B., 2009, Harmonics of circadian gene transcription in mammals, *PLoS genetics*, 5(4), p. e1000442.
- Hughes, M.E., Grant, G.R., Paquin, C., Qian, J. & Nitabach, M.N., 2012, Deep sequencing the circadian and diurnal transcriptome of Drosophila brain, *Genome research*, 22(7), pp. 1266-81.
- Hughes, M.E., Hogenesch, J.B. & Kornacker, K., 2010, JTK\_CYCLE: an efficient nonparametric algorithm for detecting rhythmic components in genome-scale data sets, *Journal of biological rhythms*, 25(5), pp. 372-80.
- Hundley, H.A. & Bass, B.L., 2010, ADAR editing in double-stranded UTRs and other noncoding RNA sequences, *Trends in biochemical sciences*, 35(7), pp. 377-83.

- Hurley, J.M., Dasgupta, A., Emerson, J.M., Zhou, X., Ringelberg, C.S., Knabe, N., Lipzen, A.M., Lindquist, E.A., Daum, C.G., Barry, K.W., Grigoriev, I.V., Smith, K.M., Galagan, J.E., Bell-Pedersen, D., Freitag, M., Cheng, C., Loros, J.J. & Dunlap, J.C., 2014, Analysis of clock-regulated genes in Neurospora reveals widespread posttranscriptional control of metabolic potential, *Proceedings of the National Academy of Sciences of the United States of America*.
- Hutchinson, J.N., Ensminger, A.W., Clemson, C.M., Lynch, C.R., Lawrence, J.B. & Chess, A., 2007, A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains, *BMC genomics*, 8, p. 39.
- Ibata, Y., Takahashi, Y., Okamura, H., Kawakami, F., Terubayashi, H., Kubo, T. & Yanaihara, N., 1989, Vasoactive intestinal peptide (VIP)-like immunoreactive neurons located in the rat suprachiasmatic nucleus receive a direct retinal projection, *Neuroscience letters*, 97(1-2), pp. 1-5.
- Imamura, K., Imamachi, N., Akizuki, G., Kumakura, M., Kawaguchi, A., Nagata, K., Kato, A., Kawaguchi, Y., Sato, H., Yoneda, M., Kai, C., Yada, T., Suzuki, Y., Yamada, T., Ozawa, T., Kaneki, K., Inoue, T., Kobayashi, M., Kodama, T., Wada, Y., Sekimizu, K. & Akimitsu, N., 2014, Long Noncoding RNA NEAT1-Dependent SFPQ Relocation from Promoter Region to Paraspeckle Mediates IL8 Expression upon Immune Stimuli, *Molecular cell*, 53(3), pp. 393-406.
- Ito, D., Seki, M., Tsunoda, Y., Uchiyama, H. & Suzuki, N., 2011, Nuclear transport impairment of amyotrophic lateral sclerosis-linked mutations in FUS/TLS, *Annals of neurology*, 69(1), pp. 152-62.
- Izumo, M., Johnson, C.H. & Yamazaki, S., 2003, Circadian gene expression in mammalian fibroblasts revealed by real-time luminescence reporting: temperature compensation and damping, *Proceedings of the National Academy of Sciences of the United States of America*, 100(26), pp. 16089-94.
- Jalali, S., Bhartiya, D., Lalwani, M.K., Sivasubbu, S. & Scaria, V., 2013, Systematic transcriptome wide analysis of lncRNA-miRNA interactions, *PloS one*, 8(2), p. e53823.
- Jouffe, C., Cretenet, G., Symul, L., Martin, E., Atger, F., Naef, F. & Gachon, F., 2013, The circadian clock coordinates ribosome biogenesis, *PLoS biology*, 11(1), p. e1001455.

- Kalantry, S. & Magnuson, T., 2006, The Polycomb group protein EED is dispensable for the initiation of random X-chromosome inactivation, *PLoS genetics*, 2(5), p. e66.
- Kalsbeek, A., Fliers, E., Franke, A.N., Wortel, J. & Buijs, R.M., 2000, Functional connections between the suprachiasmatic nucleus and the thyroid gland as revealed by lesioning and viral tracing techniques in the rat, *Endocrinology*, 141(10), pp. 3832-41.
- Kapusta, A. & Feschotte, C., 2014, Volatile evolution of long noncoding RNA repertoires: mechanisms and biological implications, *Trends in genetics : TIG*, 30(10), pp. 439-52.
- Karakashian, M.W. & Hastings, J.W., 1962, The inhibition of a biological clock by actinomycin D, *Proceedings of the National Academy of Sciences*, 48(12), pp. 2130-7.
- Karatsoreos, I.N., Yan, L., LeSauter, J. & Silver, R., 2004, Phenotype matters: identification of light-responsive cells in the mouse suprachiasmatic nucleus, *Journal of Neuroscience*, 24(1), pp. 68-75.
- Karlsson, B., Knutsson, A. & Lindahl, B., 2001, Is there an association between shift work and having a metabolic syndrome? Results from a population based study of 27 485 people, *Occupational and environmental medicine*, 58(11), pp. 747-52.
- Kiessling, S., Beaulieu-Laroche, L., Blum, I.D., Landgraf, D., Welsh, D.K., Storch, K.F., Labrecque, N. & Cermakian, N., 2017, Enhancing circadian clock function in cancer cells inhibits tumor growth, *BMC biology*, 15(1), p. 13.
- King, D.P., Zhao, Y., Sangoram, A.M., Wilsbacher, L.D., Tanaka, M., Antoch, M.P., Steeves, T.D., Vitaterna, M.H., Kornhauser, J.M. & Lowrey, P.L., 1997, Positional cloning of the mouse circadian clockgene, *Cell*, 89(4), pp. 641-53.
- Kiss, J., Leranth, C.S. & Halasz, B., 1984, Serotoninergic endings on VIP-neurons in the suprachiasmatic nucleus and on ACTH-neurons in the arcuate nucleus of the rat hypothalamus. A combination of high resolution autoradiography and electron microscopic immunocytochemistry, *Neuroscience letters*, 44(2), pp. 119-24.
- Knott, G.J., Bond, C.S. & Fox, A.H., 2016, The DBHS proteins SFPQ, NONO and PSPC1: a multipurpose molecular scaffold, *Nucleic acids research*.
- Ko, C.H. & Takahashi, J.S., 2006, Molecular components of the mammalian circadian clock, *Human molecular genetics*, 15 Spec No 2, pp. R271-7.

- Ko, C.H., Yamada, Y.R., Welsh, D.K., Buhr, E.D., Liu, A.C., Zhang, E.E., Ralph, M.R., Kay, S.A., Forger, D.B. & Takahashi, J.S., 2010, Emergence of noise-induced oscillations in the central circadian pacemaker, *PLoS biology*, 8(10), p. e1000513.
- Koike, N., Yoo, S.H., Huang, H.C., Kumar, V., Lee, C., Kim, T.K. & Takahashi, J.S., 2012, Transcriptional Architecture and Chromatin Landscape of the Core Circadian Clock in Mammals, *Science (New York, N.Y.)*, 338(6105), pp. 349-54.
- Konopka, R.J. & Benzer, S., 1971, Clock mutants of Drosophila melanogaster, *Proceedings of the National Academy of Sciences*, 68(9), pp. 2112-6.
- Kuhlman, S.J., Silver, R., Le Sauter, J., Bult-Ito, A. & McMahon, D.G., 2003, Phase resetting light pulses induce Per1 and persistent spike activity in a subpopulation of biological clock neurons, *Journal of Neuroscience*, 23(4), pp. 1441-50.
- Kume, K., Zylka, M.J., Sriram, S., Shearman, L.P., Weaver, D.R., Jin, X., Maywood, E.S., Hastings, M.H. & Reppert, S.M., 1999, mCRY1 and mCRY2 are essential components of the negative limb of the circadian clock feedback loop, *Cell*, 98(2), pp. 193-205.
- Lee, C., Etchegaray, J.P., Cagampang, F.R., Loudon, A.S. & Reppert, S.M., 2001, Posttranslational mechanisms regulate the mammalian circadian clock, *Cell*, 107(7), pp. 855-67.
- Lee, K., Loros, J.J. & Dunlap, J.C., 2000, Interconnected Feedback Loops in the Neurospora Circadian System, *Science (New York, N.Y.)*, 289(5476), pp. 107-10.
- Lee, M., Sadowska, A., Bekere, I., Ho, D., Gully, B.S., Lu, Y., Iyer, K.S., Trewhella, J., Fox, A.H. & Bond, C.S., 2015, The structure of human SFPQ reveals a coiled-coil mediated polymer essential for functional aggregation in gene regulation, *Nucleic acids research*, 43(7), pp. 3826-40.
- Lehman, M.N., Silver, R., Gladstone, W.R., Kahn, R.M., Gibson, M. & Bittman, E.L., 1987, Circadian rhythmicity restored by neural transplant. Immunocytochemical characterization of the graft and its integration with the host brain, *Journal of Neuroscience*, 7(6), pp. 1626-38.
- Lehnert, S.A., Reverter, A., Byrne, K.A., Wang, Y., Nattrass, G.S., Hudson, N.J. & Greenwood, P.L., 2007, Gene expression studies of developing bovine longissimus muscle from two different beef cattle breeds, *BMC developmental biology*, 7, p. 95.

- Le Martelot, G., Canella, D., Symul, L., Migliavacca, E., Gilardi, F., Liechti, R., Martin, O., Harshman, K., Delorenzi, M., Desvergne, B., Herr, W., Deplancke, B., Schibler, U., Rougemont, J., Guex, N., Hernandez, N., Naef, F. & the CycliX consortium, 2012, Genome-Wide RNA Polymerase II Profiles and RNA Accumulation Reveal Kinetics of Transcription and Associated Epigenetic Changes During Diurnal Cycles, *PLoS biology*, 10(11), p. e1001442.
- Li, X., Shi, L., Zhang, K., Wei, W., Liu, Q., Mao, F., Li, J., Cai, W., Chen, H., Teng, H., Li, J. & Sun, Z., 2017, CirGRDB: a database for the genome-wide deciphering circadian genes and regulators, *Nucleic acids research*.
- Liscovitch, N., Bazak, L., Levanon, E.Y. & Chechik, G., 2014, Positive correlation between ADAR expression and its targets suggests a complex regulation mediated by RNA editing in the human brain, *RNA Biol*, 11(11), pp. 1447-56.
- Liu, A.C., Welsh, D.K., Ko, C.H., Tran, H.G., Zhang, E.E., Priest, A.A., Buhr, E.D., Singer, O., Meeker, K., Verma, I.M., Doyle, F.J., Takahashi, J.S. & Kay, S.A., 2007, Intercellular coupling confers robustness against mutations in the SCN circadian clock network, *Cell*, 129(3), pp. 605-16.
- Liu, C., Weaver, D.R., Strogatz, S.H. & Reppert, S.M., 1997, Cellular construction of a circadian clock: period determination in the suprachiasmatic nuclei, *Cell*, 91(6), pp. 855-60.
- Liu, G., Mattick, J.S. & Taft, R.J., 2013, A meta-analysis of the genomic and transcriptomic composition of complex life, *Cell cycle (Georgetown, Tex.)*, 12(13), pp. 2061-72.
- Liu, S., Cai, Y., Sothern, R.B., Guan, Y. & Chan, P., 2007, Chronobiological analysis of circadian patterns in transcription of seven key clock genes in six peripheral tissues in mice, *Chronobiology international*, 24(5), pp. 793-820.
- Lo, P., Wolfson, B. & Zhou, Q., 2016, Cellular, physiological and pathological aspects of the long non-coding RNA NEAT1, *Frontiers in Biology*, 11(6), pp. 413-26.
- Lokshin, M., LeSauter, J. & Silver, R., 2015, Selective distribution of retinal input to mouse SCN revealed in analysis of sagittal sections, *Journal of biological rhythms*, 30(3), pp. 251-7.
- Lourenco, G.F., Janitz, M., Huang, Y. & Halliday, G.M., 2015, Long noncoding RNAs in TDP-43 and FUS/TLS-related frontotemporal lobar degeneration (FTLD), *Neurobiology of disease*, 82, pp. 445-54.

- Lowrey, P.L. & Takahashi, J.S., 2004, Mammalian circadian biology: elucidating genomewide levels of temporal organization, *Annu Rev Genomics Hum Genet*, 5, pp. 407-41.
- Lowrey, P.L., Shimomura, K., Antoch, M.P., Yamazaki, S., Zemenides, P.D., Ralph, M.R., Menaker, M. & Takahashi, J.S., 2000, Positional syntenic cloning and functional characterization of the mammalian circadian mutation tau, *Science*, 288(5465), pp. 483-92.
- Lück, S., Thurley, K., Thaben, P.F. & Westermark, P.O., 2014, Rhythmic degradation explains and unifies circadian transcriptome and proteome data, *Cell reports*, 9(2), pp. 741-51.
- LYONS, W.R., 1958, Hormonal synergism in mammary growth, *Proceedings of the Royal* Society of London. Series B, Biological sciences, 149(936), pp. 303-25.
- Mackey, S.R., 2007, Cold Spring Harb Symp Quant Biol, *Biological Rhythms Workshop IA: molecular basis of rhythms generation.* pp. 7-19.
- Mallepell, S., Krust, A., Chambon, P. & Brisken, C., 2006, Paracrine signaling through the epithelial estrogen receptor alpha is required for proliferation and morphogenesis in the mammary gland, *Proceedings of the National Academy of Sciences of the United States of America*, 103(7), pp. 2196-201.
- Mao, Y.S., Sunwoo, H., Zhang, B. & Spector, D.L., 2011, Direct visualization of the cotranscriptional assembly of a nuclear body by noncoding RNAs, *Nature cell biology*, 13(1), pp. 95-101.
- Mauvoisin, D., Wang, J., Jouffe, C., Martin, E., Atger, F., Waridel, P., Quadroni, M., Gachon,
  F. & Naef, F., 2014, Circadian clock-dependent and -independent rhythmic proteomes implement distinct diurnal functions in mouse liver, *Proceedings of the National Academy of Sciences of the United States of America*, 111(1), pp. 167-72.
- Mayeuf-Louchart, A., Zecchin, M., Staels, B. & Duez, H., 2017, Circadian control of metabolism and pathological consequences of clock perturbations, *Biochimie*.
- Maywood, E.S., Okamura, H. & Hastings, M.H., 2002, Opposing actions of neuropeptide Y and light on the expression of circadian clock genes in the mouse suprachiasmatic nuclei, *The European journal of neuroscience*, 15(1), pp. 216-20.

- Meijer, J.H., Groos, G.A. & Rusak, B., 1986, Luminance coding in a circadian pacemaker: the suprachiasmatic nucleus of the rat and the hamster, *Brain research*, 382(1), pp. 109-18.
- Menet, J.S., Rodriguez, J., Abruzzi, K.C. & Rosbash, M., 2012, Nascent-Seq reveals novel features of mouse circadian transcriptional regulation, *eLife*, 1, p. e00011.
- Mercer, T.R., Qureshi, I.A., Gokhan, S., Dinger, M.E., Li, G., Mattick, J.S. & Mehler, M.F., 2010, Long noncoding RNAs in neuronal-glial fate specification and oligodendrocyte lineage maturation, *BMC neuroscience*, 11, p. 14.
- Metz, A., Soret, J., Vourc'h, C., Tazi, J. & Jolly, C., 2004, A key role for stress-induced satellite III transcripts in the relocalization of splicing factors into nuclear stress granules, *Journal of cell science*, 117(Pt 19), pp. 4551-8.
- Metz, R.P., Qu, X., Laffin, B., Earnest, D. & Porter, W.W., 2006, Circadian clock and cell cycle gene expression in mouse mammary epithelial cells and in the developing mouse mammary gland, *Developmental dynamics : an official publication of the American Association of Anatomists*, 235(1), pp. 263-71.
- Miller, B.H., McDearmon, E.L., Panda, S., Hayes, K.R., Zhang, J., Andrews, J.L., Antoch, M.P., Walker, J.R., Esser, K.A., Hogenesch, J.B. & Takahashi, J.S., 2007, Circadian and CLOCK-controlled regulation of the mouse transcriptome and cell proliferation, *Proceedings of the National Academy of Sciences of the United States of America*, 104(9), pp. 3342-7.
- Miyagawa, R., Tano, K., Mizuno, R., Nakamura, Y., Ijiri, K., Rakwal, R., Shibato, J., Masuo, Y., Mayeda, A., Hirose, T. & Akimitsu, N., 2012, Identification of cis- and transacting factors involved in the localization of MALAT-1 noncoding RNA to nuclear speckles, *RNA*, 18(4), pp. 738-51.
- Moore, R.Y., 1973, Retinohypothalamic projection in mammals: a comparative study, *Brain research*, 49(2), pp. 403-9.
- Moore, R.Y. & Eichler, V.B., 1972, Loss of a circadian adrenal corticosterone rhythm following suprachiasmatic lesions in the rat, *Brain research*, 42(1), pp. 201-6.
- Moore, R.Y. & Lenn, N.J., 1972, A retinohypothalamic projection in the rat, *Journal of Comparative Neurology*, 146(1), pp. 1-14.
- Morse, D., Cermakian, N., Brancorsini, S., Parvinen, M. & Sassone-Corsi, P., 2003, No circadian rhythms in testis: Period1 expression is clock independent and

developmentally regulated in the mouse, *Molecular endocrinology (Baltimore, Md.)*, 17(1), pp. 141-51.

- Morse, D., Milos, P.M., Roux, E. & Hastings, J.W., 1989, Circadian regulation of bioluminescence in Gonyaulax involves translational control, *Proceedings of the National Academy of Sciences*, 86(1), pp. 172-6.
- Naganuma, T., Nakagawa, S., Tanigawa, A., Sasaki, Y.F., Goshima, N. & Hirose, T., 2012, Alternative 3'-end processing of long noncoding RNA initiates construction of nuclear paraspeckles, *EMBO J*.
- Nagoshi, E., Saini, C., Bauer, C., Laroche, T., Naef, F. & Schibler, U., 2004, Circadian gene expression in individual fibroblasts: cell-autonomous and self-sustained oscillators pass time to daughter cells, *Cell*, 119(5), pp. 693-705.
- Nakagawa, S., Naganuma, T., Shioi, G. & Hirose, T., 2011, Paraspeckles are subpopulation-specific nuclear bodies that are not essential in mice, *The Journal of cell biology*, 193(1), pp. 31-9.
- Nakagawa, S., Shimada, M., Yanaka, K., Mito, M., Arai, T., Takahashi, E., Fujita, Y., Fujimori, T., Standaert, L., Marine, J.C. & Hirose, T., 2014, The lncRNA Neat1 is required for corpus luteum formation and the establishment of pregnancy in a subpopulation of mice, *Development (Cambridge, England)*.
- Nakamura, W., Honma, S., Shirakawa, T. & Honma, K., 2002, Clock mutation lengthens the circadian period without damping rhythms in individual SCN neurons, *Nat Neurosci*, 5(5), pp. 399-400.
- NANDI, S., 1958, Endocrine control of mammarygland development and function in the C3H/ He Crgl mouse, *Journal of the National Cancer Institute*, 21(6), pp. 1039-63.
- Nishimoto, Y., Nakagawa, S., Hirose, T., Okano, H.J., Takao, M., Shibata, S., Suyama, S., Kuwako, K., Imai, T., Murayama, S., Suzuki, N. & Okano, H., 2013, The long noncoding RNA nuclear-enriched abundant transcript 1\_2 induces paraspeckle formation in the motor neuron during the early phase of amyotrophic lateral sclerosis, *Mol Brain*, 6(1), p. 31.
- Nussey, S. & Whitehead, S., 2001, *Endocrinology: An Integrated Approach*, BIOS Scientific Publishers, Oxford.
- Okamura, H., Miyake, S., Sumi, Y., Yamaguchi, S., Yasui, A., Muijtjens, M., Hoeijmakers, J.H. & van der Horst, G.T., 1999, Photic induction of mPer1 and mPer2 in crydeficient mice lacking a biological clock, *Science*, 286(5449), pp. 2531-4.

- Okamura, H., Yamaguchi, S. & Yagita, K., 2002, Molecular machinery of the circadian clock in mammals, *Cell and tissue research*, 309(1), pp. 47-56.
- Ooi, G.T., Tawadros, N. & Escalona, R.M., 2004, Pituitary cell lines and their endocrine applications, *Mol Cell Endocrinol*, 228(1-2), pp. 1-21.
- Panda, S., Antoch, M.P., Miller, B.H., Su, A.I., Schook, A.B., Straume, M., Schultz, P.G., Kay, S.A., Takahashi, J.S. & Hogenesch, J.B., 2002, Coordinated transcription of key pathways in the mouse by the circadian clock, *Cell*, 109(3), pp. 307-20.
- Passon, D.M., Lee, M., Rackham, O., Stanley, W.A., Sadowska, A., Filipovska, A., Fox, A.H.
  & Bond, C.S., 2012, Structure of the heterodimer of human NONO and paraspeckle protein component 1 and analysis of its role in subnuclear body formation, *Proceedings of the National Academy of Sciences of the United States of America*.
- Petrillo, E., Sanchez, S.E., Kornblihtt, A.R. & Yanovsky, M.J., 2011, Alternative splicing adds a new loop to the circadian clock, *Commun Integr Biol*, 4(3), pp. 284-6.
- Pfaffl, M.W., 2001, A new mathematical model for relative quantification in real-time RT-PCR, *Nucleic acids research*, 29(9), p. e45.
- Prasanth, K.V., Prasanth, S.G., Xuan, Z., Hearn, S., Freier, S.M., Bennett, C.F., Zhang, M.Q.
  & Spector, D.L., 2005, Regulating gene expression through RNA nuclear retention, *Cell*, 123(2), pp. 249-63.
- Preitner, N., Damiola, F., Zakany, J., Duboule, D., Albrecht, U. & Schibler, U., 2002, The orphan nuclear receptor REV-ERBα controls circadian transcription within the positive limb of the mammalian circadian oscillator, *Cell*, 110(2), pp. 251-60.
- Prosser, R.A., Dean, R.R., Edgar, D.M., Heller, H.C. & Miller, J.D., 1993, Serotonin and the mammalian circadian system: I. In vitro phase shifts by serotonergic agonists and antagonists, *Journal of biological rhythms*, 8(1), pp. 1-16.
- Ralph, M.R. & Menaker, M., 1988, A mutation of the circadian system in golden hamsters, *Science*, 241(4870), p. 1225.
- Ralph, M.R., Foster, R.G., Davis, F.C. & Menaker, M., 1990, Transplanted suprachiasmatic nucleus determines circadian period, *Science*, 247(4945), p. 975.
- Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A. & Zhang, F., 2013, Genome engineering using the CRISPR-Cas9 system, *Nature protocols*, 8(11), pp. 2281-308.

- Rea, M.A., 1998, Photic entrainment of circadian rhythms in rodents, *Chronobiology international*, 15(5), pp. 395-423.
- Reddy, A.B., Karp, N.A., Maywood, E.S., Sage, E.A., Deery, M., O'Neill, J.S., Wong, G.K., Chesham, J., Odell, M., Lilley, K.S., Kyriacou, C.P. & Hastings, M.H., 2006, Circadian orchestration of the hepatic proteome, *Current biology : CB*, 16(11), pp. 1107-15.
- Redrup, L., Branco, M.R., Perdeaux, E.R., Krueger, C., Lewis, A., Santos, F., Nagano, T., Cobb, B.S., Fraser, P. & Reik, W., 2009, The long noncoding RNA Kcnq1ot1 organises a lineage-specific nuclear domain for epigenetic gene silencing, *Development (Cambridge, England)*, 136(4), pp. 525-30.
- Rinn, J.L., Kertesz, M., Wang, J.K., Squazzo, S.L., Xu, X., Brugmann, S.A., Goodnough, L.H., Helms, J.A., Farnham, P.J. & Segal, E., 2007, Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs, *Cell*, 129(7), pp. 1311-23.
- Ripperger, J.A. & Brown, S.A., 2010, Transcriptional regulation of circadian clocks, *The Circadian Clock*, pp. 37-78.
- Saha, S., Murthy, S. & Rangarajan, P.N., 2006, Identification and characterization of a virus-inducible non-coding RNA in mouse brain, *J Gen Virol*, 87(Pt 7), pp. 19915.
- Sasaki, Y.T., Ideue, T., Sano, M., Mituyama, T. & Hirose, T., 2009, MENepsilon/beta noncoding RNAs are essential for structural integrity of nuclear paraspeckles, *Proceedings of the National Academy of Sciences of the United States of America*, 106(8), pp. 2525-30.
- Sato, T.K., Panda, S., Miraglia, L.J., Reyes, T.M., Rudic, R.D., McNamara, P., Naik, K.A., FitzGerald, G.A., Kay, S.A. & Hogenesch, J.B., 2004, A functional genomics strategy reveals Rora as a component of the mammalian circadian clock, *Neuron*, 43(4), pp. 527-37.
- Sato, T.K., Yamada, R.G., Ukai, H., Baggs, J.E., Miraglia, L.J., Kobayashi, T.J., Welsh, D.K., Kay, S.A., Ueda, H.R. & Hogenesch, J.B., 2006, Feedback repression is required for mammalian circadian clock function, *Nature genetics*, 38(3), p. 312.
- Sawaki, Y., Nihonmatsu, I. & Kawamura, H., 1984, Transplantation of the neonatal suprachiasmatic nuclei into rats with complete bilateral suprachiasmatic lesions, *Neuroscience research*, 1(1), pp. 67-72.

- Schurov, I.L., McNulty, S., Best, J.D., Sloper, P.J. & Hastings, M.H., 1999, Glutamatergic induction of CREB phosphorylation and Fos expression in primary cultures of the suprachiasmatic hypothalamus in vitro is mediated by co-ordinate activity of NMDA and non-NMDA receptors, *Journal of neuroendocrinology*, 11(1), pp. 43-51.
- Scotto-Lavino, E., Du, G. & Frohman, M.A., 2006, 3' end cDNA amplification using classic RACE, *Nature protocols*, 1(6), pp. 2742-5.
- Shearman, L.P., Sriram, S., Weaver, D.R., Maywood, E.S., Chaves, I.S., Zheng, B., Kume, K., Lee, C.C., Hastings, M.H. & Reppert, S.M., 2000, Interacting molecular loops in the mammalian circadian clock, *Science*, 288(5468), pp. 1013-9.
- Shelkovnikova, T.A., Robinson, H.K., Troakes, C., Ninkina, N. & Buchman, V.L., 2014, Compromised paraspeckle formation as a pathogenic factor in FUSopathies, *Human molecular genetics*, 23(9), pp. 2298-312.
- Shostak, A., 2017, Circadian Clock, Cell Division, and Cancer: From Molecules to Organism, *Int J Mol Sci*, 18(4).
- Silver, R., LeSauter, J., Tresco, P.A. & Lehman, M.N., 1996, A diffusible coupling signal from the transplanted suprachiasmatic nucleus controlling circadian locomotor rhythms, *Nature*, 382(6594), pp. 810-3.
- Simon, M.D., 2013, Capture hybridization analysis of RNA targets (CHART), *Curr Protoc Mol Biol*, Chapter 21, pp. Unit 21.25..
- Simon, M.D., Wang, C.I., Kharchenko, P.V., West, J.A., Chapman, B.A., Alekseyenko, A.A., Borowsky, M.L., Kuroda, M.I. & Kingston, R.E., 2011, The genomic binding sites of a noncoding RNA, *Proceedings of the National Academy of Sciences of the United States of America*, 108(51), pp. 20497-502.
- Sitzmann, B.D., Lemos, D.R., Ottinger, M.A. & Urbanski, H.F., 2010, Effects of age on clock gene expression in the rhesus macaque pituitary gland, *Neurobiology of aging*, 31(4), pp. 696-705.
- Smith, S.M. & Vale, W.W., 2006, The role of the hypothalamic-pituitary-adrenal axis in neuroendocrine responses to stress, *Dialogues in clinical neuroscience*, 8(4), pp. 383-95.
- So, W.V. & Rosbash, M., 1997, Post-transcriptional regulation contributes to Drosophila clock gene mRNA cycling, *EMBO J*, 16(23), pp. 7146-55.

- Song, Y., Yao, X. & Ying, H., 2011, Thyroid hormone action in metabolic regulation, *Protein & cell*, 2(5), pp. 358-68.
- Souquere, S., Beauclair, G., Harper, F., Fox, A. & Pierron, G., 2010, Highly ordered spatial organization of the structural long noncoding NEAT1 RNAs within paraspeckle nuclear bodies, *Molecular biology of the cell*, 21(22), pp. 4020-7.
- Spector, D.L. & Lamond, A.I., 2011, Nuclear speckles, *Cold Spring Harbor perspectives in biology*, 3(2).
- Stadler, P.F., 2010, Brazilian Symposium on Bioinformatics, *Evolution of the long noncoding RNAs MALAT1 and MENβ/varepsilon.* pp. 1-12.
- Standaert, L., Adriaens, C., Radaelli, E., Van Keymeulen, A., Blanpain, C., Hirose, T., Nakagawa, S. & Marine, J.-C., 2014, The long noncoding RNA Neat1 is required for mammary gland development and lactation, *RNA*, 20(12), pp. 1844-9.
- Stephan, F.K. & Zucker, I., 1972, Circadian rhythms in drinking behavior and locomotor activity of rats are eliminated by hypothalamic lesions, *Proceedings of the National Academy of Sciences*, 69(6), pp. 1583-6.
- Stephan, F.K., Berkley, K.J. & Moss, R.L., 1981, Efferent connections of the rat suprachiasmatic nucleus, *Neuroscience*, 6(12), pp. 2625-41.
- Storch, K.F., Lipan, O., Leykin, I., Viswanathan, N., Davis, F.C., Wong, W.H. & Weitz, C.J., 2002, Extensive and divergent circadian gene expression in liver and heart, *Nature*, 417(6884), pp. 78-83.
- Sun, Z.S., Albrecht, U., Zhuchenko, O., Bailey, J., Eichele, G. & Lee, C.C., 1997, RIGUI, a putative mammalian ortholog of the Drosophila period gene, *Cell*, 90(6), pp. 1003-11.
- Sunwoo, H., Dinger, M.E., Wilusz, J.E., Amaral, P.P., Mattick, J.S. & Spector, D.L., 2009, MEN epsilon/beta nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles, *Genome research*, 19(3), pp. 347-59.
- Swanson, L.W. & Cowan, W.M., 1975, The efferent connections of the suprachiasmatic nucleus of the hypothalamus, *Journal of comparative neurology*, 160(1), pp. 1-12.
- Takahashi, Y., Kipnis, D.M. & Daughaday, W.H., 1968, Growth hormone secretion during sleep, *J Clin Invest*, 47(9), pp. 2079-90.

- Tanaka, M., Hayashi, S., Tamada, Y., Ikeda, T., Hisa, Y., Takamatsu, T. & Ibata, Y., 1997, Direct retinal projections to GRP neurons in the suprachiasmatic nucleus of the rat, *Neuroreport*, 8(9), pp. 2187-91.
- Tashjian, A.H., Yasumura, Y., Levine, L., Sato, G.H. & Parker, M.L., 1968, Establishment of clonal strains of rat pituitary tumor cells that secrete growth hormone, *Endocrinology*, 82(2), pp. 342-52.
- Tei, H., Okamura, H., Shigeyoshi, Y., Fukuhara, C., Ozawa, R., Hirose, M. & Sakaki, Y., 1997, Circadian oscillation of a mammalian homologue of the Drosophila period gene, *Nature*, 389(6650), pp. 512-6.
- Timmons, J.A., Larsson, O., Jansson, E., Fischer, H., Gustafsson, T., Greenhaff, P.L., Ridden, J., Rachman, J., Peyrard-Janvid, M., Wahlestedt, C. & Sundberg, C.J., 2005, Human muscle gene expression responses to endurance training provide a novel perspective on Duchenne muscular dystrophy, *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 19(7), pp. 750-60.
- Toh, K.L., Jones, C.R., He, Y., Eide, E.J., Hinz, W.A., Virshup, D.M., Ptácek, L.J. & Fu, Y.H., 2001, An hPer2 phosphorylation site mutation in familial advanced sleep phase syndrome, *Science*, 291(5506), pp. 1040-3.
- Torres, M., Becquet, D., Blanchard, M.P., Guillen, S., Boyer, B., Moreno, M., Franc, J.L. & François-Bellan, A.M., 2016, Circadian RNA expression elicited by 3'-UTR IRAluparaspeckle associated elements, *eLife*, 5.
- Torres, M., Becquet, D., Blanchard, M.P., Guillen, S., Boyer, B., Moreno, M., Franc, J.L. & François-Bellan, A.M., 2017, Paraspeckles as rhythmic nuclear mRNA anchorages responsible for circadian gene expression, *Nucleus (Austin, Tex.)*, pp. 1-6.
- Triqueneaux, G., Thenot, S., Kakizawa, T., Antoch, M.P., Safi, R., Takahashi, J.S., Delaunay, F. & Laudet, V., 2004, The orphan receptor Rev-erbα gene is a target of the circadian clock pacemaker, *Journal of molecular endocrinology*, 33(3), pp. 585-608.
- Tsai, M.C., Manor, O., Wan, Y., Mosammaparast, N., Wang, J.K., Lan, F., Shi, Y., Segal, E. & Chang, H.Y., 2010, Long noncoding RNA as modular scaffold of histone modification complexes, *Science (New York, N.Y.)*, 329(5992), pp. 689-93.

- Tsuchiya, Y., Akashi, M. & Nishida, E., 2003, Temperature compensation and temperature resetting of circadian rhythms in mammalian cultured fibroblasts, *Genes to Cells*, 8(8), pp. 713-20.
- Ueda, H.R., Chen, W., Adachi, A. & Wakamatsu, H., 2002, A transcription factor response element for gene expression during circadian night, *Nature*, 418(6897), p. 534.
- van den Pol, A.N. & Tsujimoto, K.L., 1985, Neurotransmitters of the hypothalamic suprachiasmatic nucleus: immunocytochemical analysis of 25 neuronal antigens, *Neuroscience*, 15(4), pp. 1049-86.
- Van Der Horst, G.T., Muijtjens, M., Kobayashi, K. & Takano, R., 1999, Mammalian Cry1 and Cry2 are essential for maintenance of circadian rhythms, *Nature*, 398(6728), p. 627.
- Vindlacheruvu, R.R., Ebling, F.J., Maywood, E.S. & Hastings, M.H., 1992, Blockade of Glutamatergic Neurotransmission in the Suprachiasmatic Nucleus Prevents Cellular and Behavioural Responses of the Circadian System to Light, *The European journal of neuroscience*, 4(7), pp. 673-9.
- Visa, N., Puvion-Dutilleul, F., Harper, F., Bachellerie, J.P. & Puvion, E., 1993, Intranuclear distribution of poly(A) RNA determined by electron microscope in situ hybridization, *Experimental cell research*, 208(1), pp. 19-34.
- Vitaterna, M.H., King, D.P., Chang, A.-M., Kornhauser, J.M., Lowrey, P.L., McDonald, J.D., Dove, W.F., Pinto, L.H., Turek, F.W. & Takahashi, J.S., 1994, Mutagenesis and mapping of a mouse gene, Clock, essential for circadian behavior, *Science (New York, NY)*, 264(5159), p. 719.
- Vitaterna, M.H., Selby, C.P., Todo, T., Niwa, H., Thompson, C., Fruechte, E.M., Hitomi, K., Thresher, R.J., Ishikawa, T., Miyazaki, J., Takahashi, J.S. & Sancar, A., 1999, Differential regulation of mammalian period genes and circadian rhythmicity by cryptochromes 1 and 2, *Proceedings of the National Academy of Sciences of the United States of America*, 96(21), pp. 12114-9.
- Vollmers, C., Gill, S., DiTacchio, L., Pulivarthy, S.R., Le, H.D. & Panda, S., 2009, Time of feeding and the intrinsic circadian clock drive rhythms in hepatic gene expression, *Proceedings of the National Academy of Sciences*, 106(50), pp. 21453-8.

- Vollmers, C., Schmitz, R.J., Nathanson, J., Yeo, G., Ecker, J.R. & Panda, S., 2012, Circadian Oscillations of Protein-Coding and Regulatory RNAs in a Highly Dynamic Mammalian Liver Epigenome, *Cell metabolism*, 16(6), pp. 833-45.
- Watts, A.G. & Swanson, L.W., 1987, Efferent projections of the suprachiasmatic nucleus:
  II. Studies using retrograde transport of fluorescent dyes and simultaneous peptide immunohistochemistry in the rat, *Journal of Comparative Neurology*, 258(2), pp. 230-52.
- Watts, A.G., Swanson, L.W. & Sanchez-Watts, G., 1987, Efferent projections of the suprachiasmatic nucleus: I. Studies using anterograde transport of Phaseolus vulgaris leucoagglutinin in the rat, *Journal of Comparative Neurology*, 258(2), pp. 204-29.
- Webb, A.B., Angelo, N., Huettner, J.E. & Herzog, E.D., 2009, Intrinsic, nondeterministic circadian rhythm generation in identified mammalian neurons, *Proceedings of the National Academy of Sciences of the United States of America*, 106(38), pp. 16493-8.
- Wegrzyn, L.R., Tamimi, R.M., Rosner, B.A., Brown, S.B., Stevens, R.G., Eliassen, A.H., Laden, F., Willett, W.C., Hankinson, S.E. & Schernhammer, E.S., 2017, Rotating Night-Shift Work and the Risk of Breast Cancer in the Nurses' Health Studies, *American journal of epidemiology*, 186(5), pp. 532-40.
- Welsh, D.K., Logothetis, D.E., Meister, M. & Reppert, S.M., 1995, Individual neurons dissociated from rat suprachiasmatic nucleus express independently phased circadian firing rhythms, *Neuron*, 14(4), pp. 697-706.
- Welsh, D.K., Yoo, S.H., Liu, A.C., Takahashi, J.S. & Kay, S.A., 2004, Bioluminescence imaging of individual fibroblasts reveals persistent, independently phased circadian rhythms of clock gene expression, *Current biology : CB*, 14(24), pp. 2289-95.
- West, J.A., Davis, C.P., Sunwoo, H., Simon, M.D., Sadreyev, R.I., Wang, P.I., Tolstorukov, M.Y. & Kingston, R.E., 2014, The Long Noncoding RNAs NEAT1 and MALAT1 Bind Active Chromatin Sites, *Molecular cell*.
- West, J.A., Mito, M., Kurosaka, S., Takumi, T., Tanegashima, C., Chujo, T., Yanaka, K., Kingston, R.E., Hirose, T., Bond, C., Fox, A. & Nakagawa, S., 2016, Structural, super-resolution microscopy analysis of paraspeckle nuclear body organization, *The Journal of cell biology*.

- Xu, Y., Padiath, Q.S., Shapiro, R.E. & Jones, C.R., 2005, Functional consequences of a CKIdelta mutation causing familial advanced sleep phase syndrome, *Nature*, 434(7033), p. 640.
- Yagita, K., Tamanini, F., Van der Horst, G.T. & Okamura, H., 2001, Molecular mechanisms of the biological clock in cultured fibroblasts, *Science*, 292(5515), pp. 278-81.
- Yamamoto, T., Nakahata, Y., Soma, H., Akashi, M., Mamine, T. & Takumi, T., 2004, Transcriptional oscillation of canonical clock genes in mouse peripheral tissues, *BMC Mol Biol*, 5, p. 18.
- Yamazaki, S., Numano, R., Abe, M., Hida, A., Takahashi, R.-I., Ueda, M., Block, G.D., Sakaki,
  Y., Menaker, M. & Tei, H., 2000, Resetting central and peripheral circadian oscillators in transgenic rats, *Science*, 288(5466), pp. 682-5.
- Yan, L. & Okamura, H., 2002, Gradients in the circadian expression of Per1 and Per2 genes in the rat suprachiasmatic nucleus, *European Journal of Neuroscience*, 15(7), pp. 1153-62.
- Yan, L. & Silver, R., 2004, Resetting the brain clock: time course and localization of mPER1 and mPER2 protein expression in suprachiasmatic nuclei during phase shifts, *The European journal of neuroscience*, 19(4), pp. 1105-9.
- Yin, L., Wu, N., Curtin, J.C., Qatanani, M., Szwergold, N.R., Reid, R.A., Waitt, G.M., Parks, D.J., Pearce, K.H., Wisely, G.B. & Lazar, M.A., 2007, Rev-erbalpha, a heme sensor that coordinates metabolic and circadian pathways, *Science (New York, N.Y.)*, 318(5857), pp. 1786-9.
- Yoo, S.H., Yamazaki, S., Lowrey, P.L., Shimomura, K., Ko, C.H., Buhr, E.D., Siepka, S.M., Hong, H.K., Oh, W.J., Yoo, O.J., Menaker, M. & Takahashi, J.S., 2004, PERIOD2::LUCIFERASE real-time reporting of circadian dynamics reveals persistent circadian oscillations in mouse peripheral tissues, *Proceedings of the National Academy of Sciences of the United States of America*, 101(15), pp. 5339-46.
- Yoon, J.H., Abdelmohsen, K., Srikantan, S., Yang, X., Martindale, J.L., De, S., Huarte, M., Zhan, M., Becker, K.G. & Gorospe, M., 2012, LincRNA-p21 suppresses target mRNA translation, *Molecular cell*, 47(4), pp. 648-55.

- Zeng, C., Xu, Y., Xu, L., Yu, X., Cheng, J., Yang, L., Chen, S. & Li, Y., 2014, Inhibition of long non-coding RNA NEAT1 impairs myeloid differentiation in acute promyelocytic leukemia cells, *BMC cancer*, 14(1), p. 693.
- Zhang, Q., Chen, C.Y., Yedavalli, V.S. & Jeang, K.T., 2013, NEAT1 Long Noncoding RNA and Paraspeckle Bodies Modulate HIV-1 Posttranscriptional Expression, *MBio*, 4(1).
- Zhang, R., Lahens, N.F., Ballance, H.I., Hughes, M.E. & Hogenesch, J.B., 2014, A circadian gene expression atlas in mammals: implications for biology and medicine, *Proceedings of the National Academy of Sciences of the United States of America*, 111(45), pp. 16219-24.
- Zhang, Z. & Carmichael, G.G., 2001, The fate of dsRNA in the nucleus: a p54(nrb)containing complex mediates the nuclear retention of promiscuously A-to-I edited RNAs, *Cell*, 106(4), pp. 465-75.
- Zheng, B., Larkin, D.W., Albrecht, U., Sun, Z.S., Sage, M., Eichele, G., Lee, C.C. & Bradley, A., 1999, The mPer2 gene encodes a functional component of the mammalian circadian clock, *Nature*, 400(6740), pp. 169-73.