

microARN, ARN C/D et empreinte génomique parentale

CAVAILLE Jérôme

Laboratoire de Biologie Moléculaire des Eucaryotes LBME - CNRS UMR 5099 - IFR 109, Université Paul Sabatier, 118, Route de Narbonne 31062 - Toulouse Cedex, France

Auteur correspondant : CAVAILLE Jérôme, cavaille@ibcg.biotoul.fr

Reçu le 12 Mars 2007

Résumé – A travers des approches *in silico* et expérimentales, nous avons récemment identifié des centaines de petits ARN non-codants, ARNnc - ARN C/D et microARN - dont les gènes sont soumis à l'empreinte génomique parentale (EGP). Ils sont regroupés au sein de deux loci chromosomiques conservés : le domaine DLK1-GTL2 (locus Calypige, position chromosomique humaine 14q32) et le domaine SNURF-SNRPN (locus Prader-Willi, position chromosomique humaine 15q11q13). Dans cet article, l'organisation génique, le mode d'expression et les fonctions de ces petits ARN sont présentés.

Mots clés : ARN C/D / microARN / syndrome de Prader-Willi / régulation épigénétique / empreinte génomique parentale

Abstract – Micro RNA, C/D RNA and parental genoming imprinting.

Experimental and computer-assisted approaches have led us to identify hundreds of small non-coding RNA (ncRNAs) whose genes - clustered at two chromosomal domains (the DLK1-GTL2 region/human 15q11q13 and the SNURF-SNRPN/human 14q32 loc) – are subjected to genomic imprinting. Here, we discuss their genomic organization, their expression pattern and their potential functions.

Key words: C/D RNAs / microRNA / Prader-Willi / epigenetic regulation / genomic imprinting

Introduction

Le séquençage du génome humain, ainsi que celui d'autres organismes modèles, a révélé que nos chromosomes portent un nombre relativement limité de gènes - estimé à environ 25-30 000. A titre de comparaison le génome de la mouche du vinaigre *D. melanogaster* et celui du nématode *C. elegans* possèdent respectivement 13 600 et 18 424 gènes. Bien que l'épissage alternatif soit bien évidemment source d'une grande diversité d'ARNm et de protéines à partir d'un seul gène, ces informations de séquence laissent penser néanmoins que « la bio-complexité », pour peu que l'on puisse encore la définir, ne s'accompagne pas nécessairement d'un nombre croissant de gènes mais résulte probablement d'un **contrôle combinatoire** de leur expression dans le temps et dans l'espace.

Un autre fait qu'il s'agit aussi de garder à l'esprit concerne la partie codante du génome humain - c'est-à-dire les séquences exoniques traduites en

protéines - qui ne représente pas plus de 2 % de la totalité de l'ADN portée par nos chromosomes. Si pendant des années, les 98 % de la partie non-codante des génomes n'a reçu qu'un intérêt limité (on parle alors de « *junk DNA* » ou ADN rebut), aujourd'hui nombreuses sont les études - regroupées sous le terme générique « **d'ARNomique** » - qui mettent en lumière qu'une fraction importante des génomes de mammifère est transcrite et génère des ARN dit « **non-codants (ARNnc)** » (Mattick & Makunin, 2006 ; Carninci & Hayashizaki, 2007). La complexité du répertoire de ces ARNnc égale vraisemblablement celle des gènes codant pour des protéines et leur compréhension est littéralement en passe de révolutionner notre façon d'appréhender **l'évolution, l'organisation et l'expression des génomes eucaryotes**. Parmi les ARNnc les mieux caractérisés à ce jour, on trouve les petits ARN impliqués dans les voies de **l'interférence ARN** (ou ARNi) tels que les **miARN, siARN, rasiARN** et autres **piARN**,

requis pour la mise en silence de gènes à travers une étonnante palette de mécanismes moléculaires : ils dirigent des clivages dans des ARNm, bloquent la traduction des ARNm, voire modulent la structure de la chromatine (Kim, 2006 ; Kloosterman & Plasterk, 2006).

Un autre niveau de complexité, pas directement apparent à la simple lecture des séquences génomiques, concerne les **réseaux de régulations épigénétiques** capables de modifier l'activité des gènes de manière héritable à travers la mitose (voire la méiose) et ce, indépendamment de toutes les altérations génétiques inscrites dans l'ADN. Ces phénomènes jouent un rôle critique au cours du développement (e.g. la plupart de nos cellules ont le même génome et pourtant un neurone n'exprime pas le même jeu de gènes qu'une cellule musculaire ou une cellule de peau...), du maintien de la structure des chromosomes (centromères, télomères), de l'empreinte génomique parentale (cf. ci-dessous), de l'inactivation du chromosome X chez les femelles de mammifères, de la totipotence des cellules souches, des processus de différenciation cellulaire... (Bernstein *et al.*, 2007 ; Surani *et al.*, 2007). Au niveau moléculaire, à travers ce terme de « régulations épigénétiques », sont regroupés les différents travaux visant à comprendre les multiples états héréditaires de la structure de la chromatine et l'arsenal des modifications chimiques capables de les moduler telles que la méthylation de l'ADN (méthylation sur les cytosines) ou les modifications post-traductionnelles des queues N^{ter} des histones (méthylation, phosphorylation, ubiquitinylation). Il est en effet largement admis que de telles marques épigénétiques forment « un code » ou « une mémoire », qui en modifiant l'ouverture de la fibre de la chromatine et/ou la position relative d'un locus au sein du noyau, module l'**accessibilité de la molécule d'ADN** vis-à-vis des facteurs liant l'ADN (facteurs de transcription et autres facteurs liant la chromatine) permettant ainsi d'enclencher (ou au contraire de restreindre) l'expression d'un gène et/ou d'un domaine chromosomique (Bernstein *et al.*, 2007 ; Kouzarides, 2007). Ces modifications sont généralement **réversibles**, et bien qu'héréditaires après chaque division cellulaire, certaines d'entre elles sont relativement **labiles**, avec notamment une mise en place parfois stochastique et un maintien susceptible d'être modulé par l'environnement (Feil, 2006).

Empreinte génomique parentale : 1 + 1 = 1. Tout individu hérite de deux lots de chromosomes homologues (à l'exception du X et Y chez les mâles) provenant de sa mère et de son père. Conformément aux lois classiques de l'hérédité, la grande majorité des allèles parentaux s'exprime indépendamment de l'origine maternelle ou paternelle du chromosome qui les porte. Certains gènes - les gènes soumis à l'**em-**

preinte génomique parentale (l'EGP) - défient les lois mendéliennes dans la mesure où un seul des deux allèles - hérité de l'un ou de l'autre des parents - est fonctionnel (Reik and Walter, 2001). La plupart des gènes soumis à l'empreinte sont regroupés au sein de **grandes régions chromosomiques** (ou *clusters*) dont la taille peut atteindre quelques mégabases. Ces *clusters* hébergent à la fois des gènes exprimés à partir du chromosome d'origine paternelle ou maternelle ainsi que parfois des gènes qui échappent complètement à ces régulations. Chaque *cluster* est caractérisé par un « **centre d'empreinte** » (**CE**), élément en *cis* complexe, qui chevauche généralement des îlots CpG dont le degré de méthylation sur les cytosines diffère considérablement selon son passage par la lignée germinale mâle ou femelle. On parle de régions différenciellement méthylées (ou RDM). Ces méthylations sont **réversibles** puisqu'un chromosome d'origine maternelle chez un individu de sexe masculin sera transmis à sa progéniture avec une empreinte de type paternel (et vice versa pour les individus de sexe féminin). Par des mécanismes qui demeurent encore peu compris, ces CE fonctionnent différemment sur les deux chromosomes et régissent, en *cis* et à très longue distance, l'expression coordonnée de tous les gènes d'un même *cluster*. Au sein des loci soumis à l'EGP, la présence de gènes atypiques générant des **ARN non-codants** est aussi un thème récurrent (Sleutels *et al.*, 2002 ; Bernstein and Allis, 2005 ; O'Neill, 2005 ; Mancini-Dinardo *et al.*, 2006 ; Yang and Kuroda, 2007). Ils sont généralement transcrits dans la direction opposée à celui des gènes adjacents avec le plus souvent une **empreinte réciproque** par rapport à ces derniers (e.g. si les ARNnc sont exprimés à partir de l'allèle maternel, alors les gènes au voisinage sont exprimés à partir de l'allèle paternel). Bien que la fonction de la plupart de ces **ARNnc antisens** reste peu documentée, il est proposé que ce type d'organisation contrôle l'expression mono-allélique de ces paires (ou de ces groupes) de gènes. Sur le plan mécanistique il n'est pas encore clair si c'est le fait de transcrire ces régions et/ou si c'est l'ARN néo-synthétisé qui est important pour la mise en silence (Sleutels *et al.*, 2002 ; Mancini-Dinardo *et al.*, 2006 ; Seidl *et al.*, 2006).

microARN : 1, 2, 3, 4, 5, 6, 7... 20 : Silence !

Les microARN représentent une classe d'ARN non-codants (<http://microrna.sanger.ac.uk/sequences/>) impliquée dans **la mise en silence de gènes** essentiellement au niveau post-transcriptionnel. Les gènes des microARN ont un patron d'expression très varié, avec pour certains d'entre eux une expression ubiquitaire ou au contraire restreinte à un type cellulaire et/ou à un stade de développement précis. Ils apparaissent néanmoins tous transcrits sous forme d'un ARN précurseur, le **pri-microARN**, qui est converti dans le noyau par le complexe Drosha/DGCR8 en un

intermédiaire d'environ 70 nt, le **pré-miARN**, capable de se replier en une structure secondaire de type « épingle à cheveux irrégulière ». Ce pré-miARN est par la suite exporté dans le cytoplasme pour y être clivé par Dicer qui libère un ARN \sim 22 nt provenant d'une des deux bras de la tige. C'est ce court ARN qui est nommé **microARN** et qui représente l'entité fonctionnelle (Zamore and Haley, 2005).

Le microARN interagit avec des protéines spécifiques pour former une ribonucléoparticule stable (appelée **RISC** ou **miRNP**) au sein de laquelle, le microARN joue le rôle de guide en s'appariant avec des ARNm-cibles (généralement dans leurs régions 3'-UTR). Cette interaction microARN : ARNm-cible provoque **une diminution de la quantité de protéines** synthétisées et/ou une **déstabilisation de l'ARNm-cible**. Les déterminants moléculaires qui engagent l'ARNm-cible vers ces deux voies semblent essentiellement dictés par le **degré de complémentarité** entre le 3'-UTR-cible et le microARN : (i) un appariement parfait induit la dégradation rapide de l'ARNm-cible avec un clivage endonucléolytique au milieu de l'hybride miARN : ARNm suivi par une rapide dégradation exonucléasique (ii) une complémentarité imparfaite (i.e. présence de nucléotides non-appariés dans le duplex) entraîne préférentiellement une diminution de la quantité de protéine produite à partir de l'ARNm-cible. Dans ce cas de figure, la région 5' du microARN (position +2-+7) est critique pour le ciblage et la spécificité de l'appariement. Cette dichotomie du mode d'action des microRNA reste toutefois à nuancer puisque des travaux indiquent que la durée de vie des ARN-cibles imparfaitement complémentaires est aussi diminuée (Bagga *et al.*, 2005; Lim *et al.*, 2005). De même, les mécanismes moléculaires qui sous-tendent le blocage de la traduction restent très controversés (Pillai, 2005; Pillai *et al.*, 2006), avec pour certains auteurs un rôle prépondérant au niveau de l'initiation ou au contraire au niveau de l'élongation voire la dégradation du polypeptide naissant (Pillai *et al.*, 2005; Nottrott *et al.*, 2006; Petersen *et al.*, 2006; Pillai *et al.*, 2006). A de rares exceptions près (Seitz *et al.*, 2003; Yekta *et al.*, 2004), la grande majorité des microARN chez les animaux possède des appariements imparfaits, rendant particulièrement difficile l'identification des ARNm-cibles. Ainsi, la compréhension de leurs fonctions physiologiques chez les mammifères est encore relativement limitée. Néanmoins, nombreux sont les travaux indiquant une contribution majeure des microARN dans la régulation de processus moléculaires et cellulaires importants, aussi bien au cours du développement qu'à l'âge adulte (Kloosterman and Plasterk, 2006). Les analyses informatiques sont aussi compatibles avec la notion d'un réseau complexe d'interac-

tions entre ARNm-cibles et microARN puisque pas moins de **10-30 % des ARNm** humains susceptibles d'être régulés par un(des) microARN (en moyenne un microARN apparaît capable d'interagir avec 100-200 ARNm-cible différentes). Ces analyses laissent entrevoir **un contrôle combinatoire** de l'expression des gènes avec notamment l'intervention de différents microARN régulant un même ARNm-cible (Bartel and Chen, 2004). Ainsi, en modulant l'expression de multiples gènes, les microARN contrôlent **l'état de prolifération** et de **différenciation cellulaire**, et par là même participe vraisemblablement à l'acquisition de l'identité cellulaire ou tissulaire. Enfin, les microARN pourraient aussi constituer une source de **variabilité phénotypique** entre espèces relativement proches sur le plan phylogénétique voire entre individus d'une même espèce (Niwa and Slack, 2007).

ARN C/D guides de méthylation en 2'-O-ribose : 1..2..3..4..5 : Méthyle! Les ARN C/D sont des ARN de courte taille (\sim 65-300 nt), métaboliquement stables et retrouvés aussi bien chez les eucaryotes que chez les archaebactéries (<http://www-snoRNA.biotoul.fr/>). Comme leurs noms l'indiquent, ces ARN sont caractérisés par la présence de deux signatures moléculaires, séquences conservées appelées boîtes C (5'-RUGAUGA-3') et D (5'-CUGA-3'), qui dans la plupart des cas sont positionnées respectivement aux extrémités 5' et 3' de l'ARN. Des motifs dégénérés de ces boîtes, nommés C' et D', sont aussi retrouvés en position interne. Les motifs C et D interagissent entre eux et forment un repliement spécifique, appelé le « *kink-turn* », véritable **plate-forme structurale** qui recrute des protéines spécifiques, notamment 15.5 kDa/snu13, puis la fibrillarine, puis nop56 et nop58, pour former une particule ribonucléoprotéique C/D (RPN C/D) qui peut s'accumuler soit dans le **nucléole** (on parle de **snoARN C/D**) soit dans **les corps de Cajal** (on parle de **scaARN**) (Bachellerie *et al.*, 2002; Kiss, 2002).

De nombreux ARN C/D possèdent **des segments antisens phylogénétiquement conservés** (10-20 nt) positionnés en aval des motifs D ou D'. Ces segments antisens interagissent, principalement *via* des appariements Watson-Crick, avec des ARN cellulaires et forment un duplex canonique qui sert « **d'ARN guides** » de **méthylation en 2'-O-ribose**. En effet, le ribose de la position appariée à la 5^{ème} base en amont de la boîte D est systématiquement 2'-O-méthylé et ce, quelle que soit la nature de la base (Cavaille *et al.*, 1996; Kiss-Laszlo *et al.*, 1996). L'activité 2'-O-ribose méthylase est portée par la fibrillarine. Les snoARN C/D modifient les ARN ribosomiques (transcrits Pol-I) et le petit snARN U6 du spliceosome (transcrits Pol-III) alors que les scaARN modifient les petits snARN du spli-

ceosome U1, U2, U4 et U5 (transcrits Pol-II). Chez les archaebactéries, les ARNt sont aussi modifiés via les ARN C/D. Dans la grande majorité des cas, la fonction moléculaire des méthylations introduites reste à déterminer. Il est cependant admis qu'elles sont capables de moduler la structure secondaire et tertiaire des ARN modifiés, facilitant les interactions et autres réactions chimiques dans lesquelles elles sont impliquées (e.g. traduction et épissage des introns respectivement dans le cas du ribosome et du spliceosome). Il existe aussi des ARN C/D qui ne participent pas à ces réactions de 2'-O-méthylation mais qui, en s'appariant avec l'ARNr, joue un rôle de « **chaperons moléculaires** » pour faciliter certaines réactions de clivage qui transforment le pré-ARNr en ARNr matures. Enfin, de nombreux ARN C/D de mammifères, incluant les ARN C/D imprimés décrits ci-dessous, présentent toutes les caractéristiques des guides de méthylation mais ne présentent pas de segments de complémentarité vis-à-vis des ARNr ou des snARN du spliceosome. La fonction de ces « **ARN C/D orphelins** » demeure inconnue et la possibilité que des ARNm puissent être modifiés n'est pas formellement exclue (cf. ci-dessous).

Petits ARN et empreinte génomique parentale

En combinant approches expérimentales (séquençage systématique de banques d'ADNc) et *in silico* (recherche de motifs de séquence caractéristiques des ARN C/D), nous avons identifié des centaines de gènes codant pour des **microARN** et des **ARN C/D**, principalement regroupés dans 2 loci soumis à l'EGP : le locus **Dlk1-Gtl2** (ou locus Callipyge en position 14q32 chez l'Homme) et le locus **Snurf-Snrpn** (ou locus Prader-Willi en position 15q11q13 chez l'Homme) au sein desquels ils sont respectivement exprimés à partir de l'allèle maternel et paternel (Cavaillé *et al.*, 2000; Cavaillé *et al.*, 2001; Cavaillé *et al.*, 2002; Seitz *et al.*, 2003; Seitz *et al.*, 2004a; Davis *et al.*, 2005). L'organisation et l'expression des gènes de ces petits ARN présentent un certain nombre de caractéristiques communes (Cf. Figure 1A). Ces gènes : (i) sont pour la plupart répétés et regroupés en *clusters* de gènes dont les séquences partagent de fortes homologies (ce type d'organisation répétée est conservée chez les mammifères) (ii) sont pour la plupart d'entre eux localisés dans des introns d'autres gènes eux-même ne possédant pas de cadre ouvert de lecture évident (iii) ont un patron d'expression spatio-temporel complexe avec une expression marquée dans le cerveau adulte ; (iv) sont relativement peu conservés à travers l'évolution puisque présents uniquement chez les mammifères placentaires (euthériens) ; (iv) sont

pour la plupart « orphelins » (pas de segments ARN antisens).

Le domaine SNURF-SNRPN

La région chromosomique humaine 15q11q13 (Chromosome 7C murin) soumis à l'EGP contient au moins 7 gènes codant pour des protéines : UBE3A et ATP10C exprimés à partir de l'allèle maternel et NDN, MAGEL2, MKRN3, SNURF-SNRPN exprimés à partir du chromosome paternel. Ce locus est caractérisé par une grande région non-codante positionnée entre les gènes SNURF-SNRP et UBE3A et contenant de multiples gènes d'ARN C/D (HBII-13, HBII-436, HBII-437, HBII-438a, HBII-438b, HBII-52 et HBII-85), avec HBII-85 et HBII-52 codés par des gènes homologues répétés respectivement 27 et 47 fois. La plupart des gènes des ARN C/D est localisée dans des introns d'énormes gènes non-codants s'étalant sur pas moins de 460 kb et 1MB de séquence génomique, respectivement chez l'homme et la souris (Runte *et al.*, 2001; Landers *et al.*, 2004). Cette unité de transcription génère, non seulement les ARN C/D, mais aussi de multiples transcrits alternativement épissés dont la fonction reste à déterminer. L'expression du gène-hôte, et donc celle des ARN C/D, est régulée par un CE situé en amont du gène SNURF-SNRPN (méthylé sur l'allèle maternel et non-méthylé sur l'allèle paternel), région génomique qui constitue aussi vraisemblablement le site de l'initiation de la transcription du gène des ARN C/D. Fait remarquable, la partie distale du grand transcrit chevauche dans l'orientation antisens le gène UBE3A. Il est proposé que ce type d'orientation sens-antisens joue un rôle *en cis* dans la mise en silence de l'allèle UBE3A sur le chromosome paternel (Rougeulle *et al.*, 1998; Chamberlain and Brannan, 2001; Runte *et al.*, 2001), même si de récents travaux ont nuancé cette hypothèse (Le Meur *et al.*, 2005).

L'ARN HBII-85 et le syndrome de Prader-Willi- Le locus humain 15q11q13 est associé à une maladie orpheline, le **syndrome de Prader-Willi**, une maladie rare dont la fréquence d'apparition est estimée à 1/25 000-1/30 000. Ce syndrome est caractérisé par des phénotypes complexes, comme une hypotonie majeure à la naissance avec des difficultés pour le nourrisson à téter et des comportements inhabituels vis-à-vis de la nourriture qui conduit les enfants et les adultes à manger énormément (perte du sentiment de satiété). Cet attrait exagéré pour la nourriture, le plus souvent associé à des difficultés d'apprentissage, peuvent conduire à une obésité morbide (<http://www.prader-willi.fr/>). Les études actuelles indiquent que ce syndrome complexe résulte de l'absence d'un (ou de plusieurs gènes) à expression

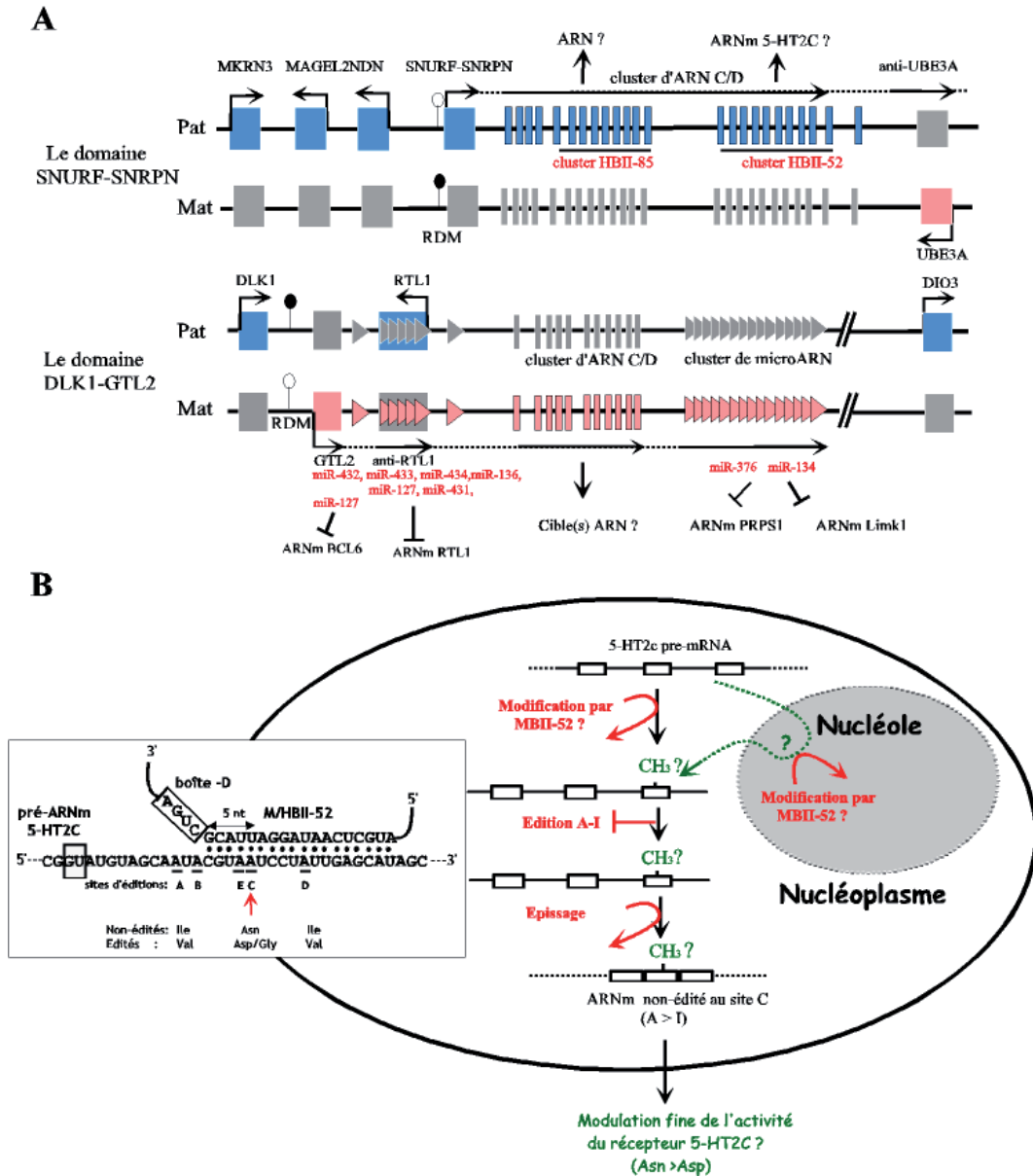


Fig. 1. A) Représentation schématique du domaine SNURF-SNRPN (15q11q13 humain; en haut) et DLK1-Gtl2 (14q32 humain; en bas) qui hébergent des gènes codant pour des protéines (rectangles) et de nombreux gènes codant pour des ARN C/D (petits rectangles verticaux) et des microARN (têtes de flèche), avec en bleu et rose, les gènes exprimés respectivement à partir de l'allèle paternel et maternel. (Les régions différemment méthylées (RDM) qui servent de « centre d'empreinte » sont représentées par des cercles pleins et vides respectivement pour signaler une hyperméthylation et une hypométhylation). Le sens de la transcription des gènes est indiqué par les flèches. Par souci de clarté les gènes hôtes hébergeant les ARN C/D et microARN ne sont pas indiqués. **B**) Modèle du mode d'action de MBII-52 : Nous proposons que l'ARN C/D M/HBII-52 puisse modifier l'ARNm 5-HT2C, (dans le nucléole et/ou le nucléoplasme ?) et que la 2'-O méthylation introduite inhibe sélectivement l'édition au site C, modulant ainsi de manière subtile les propriétés fonctionnelles du récepteur. Ce modèle est basé sur nos résultats (Vitali et al., 2005). Encadré : Modèle de l'appariement entre l'ARNm 5-HT2C et le petit ARN C/D MBII-52. La 5^{ème} position appariée en amont de la boîte D (flèche rouge) correspond au site d'édition C, une adénosine connue pour être déaminée en inosine. Cet événement d'édition altère l'information génétique de l'ARNm (I=G) et le changement en acide aminé dans le récepteur (Asn-Gly) change les propriétés de transduction du signal de ce dernier, notamment en réduisant sa capacité de couplage aux protéines G.

paternelle positionné(s) au locus 15q11q13 (Nicholls and Knepper, 2001). En effet, dans environ 70-75 % des cas, le syndrome de Prader-Willi résulte d'une grande délétion *de novo* sur le chromosome paternel, délétion qui englobe l'ensemble des gènes soumis à l'EGP de cet intervalle chromosomique (les autres cas résultent de matUPD15 (e.g. 2 chromosomes 15 d'origine maternelle) ou de rares translocations et/ou défaut de la mise en place de l'empreinte parentale). Si par contre cette même délétion prend place non pas sur le chromosome d'origine paternelle mais sur celui d'origine maternelle, alors les individus présentent un tout autre syndrome : le syndrome d'Angelman caractérisé par une déficience mentale sévère, une ataxie, une hyperactivité motrice et une hyperexcitabilité. L'identité du(des) gène(s) impliqué(s) dans le syndrome de Prader-Willi est inconnue. Cependant l'analyse de modèles murins visant à reproduire certains aspects de la maladie et l'identification de rares translocations chromosomiques chez l'Homme ont permis de définir un intervalle génomique minimal (localisé en aval de SNRPN et en amont du *cluster* des gènes M/HBII-52) susceptible d'héberger un(des) gènes dont des défauts d'expression contribuent de manière significative à l'apparition de la maladie (Wirth *et al.*, 2001; Gallagher *et al.*, 2002; Ding *et al.*, 2005; Schule *et al.*, 2005). A ce jour, les seuls gènes identifiés dans cette région - et conservés entre la souris et l'homme - sont les gènes répétés du petit ARN M/HBII-85, dont la fonction demeure inconnue.

L'ARN HBII-52 et le contrôle post-transcriptionnel du pré-ARNm 5HT2C - Bien que la perte d'expression de l'ARN C/D M/HBII-52 n'apparaissent pas être un déterminant critique pour l'apparition du syndrome de Prader-Willi (Burger *et al.*, 2002; Runte *et al.*, 2005), ce petit ARN reste très intéressant car il présente une longue complémentarité (18 nt) conservée avec l'ARNm codant pour un récepteur à la sérotonine, le variant 5-HT2c (Cavaille *et al.*, 2000) (Figure 1B, encadré). Ce segment antisens chevauche une région de l'ARNm 5-HT2C connue pour subir un épissage alternatif mais aussi des événements d'édition A vers I (déamination d'adénosines en inosines) sur 5 positions (sites A-E) par des enzymes spécifiques : Adar1 et Adar2 (Burns *et al.*, 1997). Ces désaminases agissent avant l'épissage (donc au niveau du pré-ARNm) car le substrat ARN reconnu par les Adars est une longue structure secondaire mettant en jeu des appariements irréguliers entre l'exon portant les sites à éditer et l'intron adjacent en 3'. C'est un phénomène d'une rare complexité puisque, selon les aires de cerveaux, l'édition au niveau de chaque site et les diverses combinaisons entre sites diffèrent considérablement (théoriquement 32 ARNm possibles avec 24 isoformes protéiques). Fait intéressant, le site prédit pour être

2'-O-méthylé par M/HBII-52 correspond précisément à un de ces sites : le site C (e.g. apparié à la 5^{ème} base en amont de la boîte D du petit ARN C/D). Durant la traduction, l'inosine est reconnue par le ribosome comme une guanosine et par conséquent, l'édition A-I a le potentiel d'altérer les capacités codantes de l'ARNm édité, notamment lorsque cette déamination touche la première ou la deuxième position informative des codons. Dans le cas de l'ARNm 5-HT2c, l'édition affecte un segment de l'ARNm codant pour la seconde boucle intracellulaire du récepteur requise pour le couplage du récepteur avec les protéines G. L'édition de l'ARNm 5-HT2c en changeant l'identité des acides aminés altère considérablement les propriétés du récepteur, avec notamment une baisse de l'activité constitutive du récepteur issu des ARNm édités (essentiellement pour ceux édités aux sites E et C) (Burns *et al.*, 1997; Niswender *et al.*, 1999; Wang *et al.*, 2000).

Nos observations suggèrent donc que l'appariement d'un ARN C/D, et donc vraisemblablement la méthylation guidée par ce dernier, pourrait réguler l'efficacité de l'édition A vers I. En effet, des évidences expérimentales *ex vivo* confortent un modèle selon lequel les ARN C/D ont le potentiel d'inhiber de manière sélective l'édition de substrats ARN modèles (Vitali *et al.*, 2005), laissant ainsi entrevoir un rôle régulateur de M/HBII-52 dans le contrôle post-transcriptionnel du pré-ARNm 5-HT2c. Démontrer formellement un tel modèle et ses conséquences fonctionnelles représente un réel défi expérimental, et ce d'autant plus qu'il n'existe pas de modèles animaux simples et de méthodologies suffisamment sensibles pour cartographier la présence de groupes méthyles sur des substrats aussi peu abondants que peuvent l'être les ARNm. Néanmoins, nous proposons que trois régulations post-transcriptionnelles catalysées par trois machineries nucléaires distinctes opèrent sur le pré-ARNm 5-HT2c : (i) la méthylation du site C via la RNP C/D MBII-52. (ii) l'édition A-I via les enzymes Adar1 et Adar2; (iii) l'épissage via les snRNPs (le spliceosome). Ces régulations sont largement antagonistes les unes des autres puisque l'épissage inhibe l'édition (car il élimine la structure secondaire reconnue par les Adars) et la méthylation en 2'-O-ribose a le potentiel de bloquer l'édition (nos travaux). La chronologie de ces événements pourrait donc être la suivante (Figure 1B) : le pré-ARNm 5-HT2C (tout ou une fraction de ce dernier?) est d'abord modifié au site C par l'ARN H/MBII-52 (dans le compartiment nucléolaire?), puis les pré-ARNm sont reconnus par les Adars, avec les formes 2'-O-méthylées résistantes à la déamination. En revanche, si la cinétique d'assemblage du spliceosome est plus rapide que celle de la reconnaissance du pré-ARNm par les Adars il y a production d'ARNm non-édités. L'appariement de

M/HBII-52 est aussi susceptible de réguler l'épissage alternatif du pré-ARNm 5-HT2c (Kishore and Stamm, 2006), bien que dans notre système expérimental aucun effet significatif sur l'épissage n'ait été mesuré à ce jour (nos données non publiées). Cette interaction de H/MBII-52 C/D avec l'ARN 5-HT2C représente le premier cas suggérant fortement l'intervention d'un ARN C/D dans la régulation du métabolisme d'un (pré-)ARNm, ouvrant ainsi de nombreuses perspectives encore inexplorées quant aux fonctions associées aux ARN C/D.

Le locus *Dlk1-GTL2*

Le domaine *DLK1-GTL2*, d'une taille d'environ 1 Mbase, est positionné au niveau de l'intervalle chromosomique humain 14q32 (correspondant à la partie distale du chromosome 12 murin). Il contient 3 gènes exprimés uniquement à partir de l'allèle paternel (*DLK1*, *RTL1* et *DIO3*) et de multiples gènes non-codant s'exprimant uniquement à partir de l'allèle maternel : (i) *Gtl2*, générant de multiples transcrits polyadénylés et alternativement épissés, (ii) un ARN antisens à *Rtl1* (iii) de nombreux ARN C/D ($n \sim 40$) et microARN ($n \sim 50$) qui, comme au locus Prader-Willi décrit précédemment, sont souvent imbriqués dans des introns d'autres gènes non-codants (respectivement *Rian* et *Mirg* chez la souris). Étrangement, le locus *Dlk1-Gtl2* de rat (6q32) contient une grande unité de transcription : le cluster *Bsr* avec 86 copies du gène de l'ARN C/D *RBII-36* fortement homologues et répétées en tandem sur environ 100 kb (Cavaille *et al.*, 2001 et résultats non publiés). Curieusement, ce gène n'est pas retrouvé ni dans le génome humain, ni dans celui de la souris indiquant une amplification génique de ce locus qui a vraisemblablement pris place après la divergence rat-souris (données non publiées). L'expression mono-allélique de l'ensemble du locus est gouvernée par un CE, appelé IG-DMR, localisé entre les gènes *Dlk1* et *Gtl2* (Lin *et al.*, 2003), dont les allèles paternels et maternels sont respectivement hyper- et hypométhylés. Bien que la synthèse de ces ARNnc reste encore mal compris, il reste envisageable que tous ces ARN dérivent, après maturation post-transcriptionnelle, d'un unique transcrit dont le gène serait initié au niveau du promoteur *Gtl2* (Seitz *et al.*, 2004a; Seitz *et al.*, 2004b).

microARN, mise en silence du gène *Rtl1* et le conflit d'intérêt sexuel - Au locus *Dlk1-Gtl2* certains microRNAs exprimés à partir de l'allèle maternel (miR-127, miR-136, miR-431, miR-432, miR-433 et miR-434) dérivent d'un **ARN antisens naturel** du gène *Rtl1* dont l'expression est restreinte au chromosome paternel (Seitz *et al.*, 2003). *Rtl1*, de fonction encore inconnue, est un gène sans intron qui est conservé chez les mammifères et

qui présente des homologies de séquences avec les domaines gag et pol des retrotransposons de la famille Sushi-ichi/Ty3 (Youngson *et al.*, 2005). De par leur mode d'expression, ces microARN sont donc pleinement complémentaires à l'ARNm *Rtl1* (Seitz *et al.*, 2003) et sont par conséquent capables d'induire des clivages spécifiques dans ce transcrit (Davis *et al.*, 2005). Avec miR-196 (Yekta *et al.*, 2004), ce sont les seuls miRNA décrits à ce jour chez les animaux qui sont capables de se comporter comme des siARN. Alors que la mise en silence de *Rtl1* via la voie de l'ARNi opère vraisemblablement en *trans* et dans le cytoplasme, un rôle *en cis* dans le noyau de la machinerie de l'ARNi pour contrôler l'expression mono-allélique du gène *Rtl1* (en ciblant la formation d'hétérochromatine au voisinage de *Rtl1* sur le chromosome maternel) est une hypothèse séduisante qui reste à tester (Lin *et al.*, 2003; Seitz *et al.*, 2003). Ce type de régulation entre gènes imprimés de manière réciproque supporte certaines théories évolutives, notamment la théorie du **conflit d'intérêt sexuel**, visant à expliquer l'émergence de l'empreinte parentale chez les mammifères placentaires (Davis *et al.*, 2005; Lewis & Redrup, 2005). Cette théorie est essentiellement fondée sur un antagonisme entre la fonction des gènes à expression paternelle qui exercent un effet positif sur la croissance embryonnaire en facilitant notamment les échanges nutritionnels à travers le placenta alors que les gènes à expression maternelle tendent au contraire à contrecarrer la fonction des gènes à expression paternelle (Moore & Haig, 1991).

microARN et hérédité non-mendélienne chez le mouton ? - Le locus *Dlk1-Gtl2* est non seulement soumis à l'EGP mais est aussi le siège d'une hérédité non-mendélienne décrit chez le mouton : la **dominance sur-polaire** (Georges *et al.*, 2003). Ce phénomène se traduit par le **phénotype Callipyge (Clpg)**, essentiellement caractérisé par une hypertrophie musculaire postnatale de l'arrière-train des animaux. La transmission de ce caractère présente l'originalité suivante : seule la classe des hétérozygotes ayant hérité la mutation *Clpg* de leur père- et non de leur mère! - développe le phénotype et de manière encore plus surprenante, les doubles homozygotes sont... sauvages. Au niveau moléculaire, la mutation *Clpg* est une simple transition A vers G, située entre les gènes *Dlk1* et *Gtl2* (Freking *et al.*, 2002; Smit *et al.*, 2003) qui active en *cis* l'expression de l'ensemble des gènes du locus sans pour autant en affecter l'expression mono-allélique (Charlier *et al.*, 2001). Par exemple, les gènes *Dlk1* et *Rtl1* sont sur-exprimés lorsque la mutation est transmise par le père alors que ce sont les gènes non-codants (*Gtl2*, ARN C/D, microARN et leur gène hôte) qui le sont lorsque cette même mutation est cette fois transmise par la mère. L'ensemble de ces observations conduit à penser que la surexpression de

Dlk1 et/ou de Rtl1 est susceptible de provoquer l'apparition du phénotype Callipyge (Davis *et al.*, 2004). Mais alors comment expliquer que les doubles homozygotes - qui génèrent de fortes quantités d'ARNm Dlk1 et Rtl1 - soient aussi phénotypiquement sauvages? Michel Georges et ses collègues ont proposé que dans le contexte Clpg/Clpg, un(des) facteur(s) en *trans* provenant du chromosome maternel puisse(nt) « antagoniser » l'effet du surdosage des gènes Dlk1 et/ou de Rtl1. Sachant le rôle négatif qu'exercent les microARN sur l'expression des gènes, ils représentent des candidats sérieux pour remplir ce rôle de répresseur, en particulier les miARN antisens capables de mettre en silence Rtl1 décrit précédemment (Georges *et al.*, 2003; Seitz *et al.*, 2004b; Davis *et al.*, 2005; Royo *et al.*, 2006)

miR-134 et la plasticité synaptique ?

miR-134, un microARN du grand cluster en amont des ARN C/D, est détecté dans les **épinés dendritiques** de neurones d'hippocampe, compartiment dans lequel il inhibe la traduction localisée de l'ARNm Limk1 dont le gène est induit par le BDNF (*Brain-derived neurotrophic factor*). L'interaction miR-134 : : Limk1 inhibe la croissance des épinés dendritiques suggérant que miR-134 puisse être impliqué dans un réseau de régulations sous-jacent à la maturation et à la plasticité synaptique (Schratt *et al.*, 2006; Tai and Schuman, 2006). De manière remarquable, un traitement des neurones avec du BDNF s'accompagne d'une levée de l'inhibition de la traduction de Limk1 via miR-134. Ces données sont compatibles avec un modèle selon lequel lors d'une stimulation synaptique, miR-134 n'étant plus capable de réprimer l'ARNm Limk1, la protéine Limk1 induit un changement dans la morphologie des épinés dendritiques. Les modalités du transport de miR-134 (du corps cellulaire vers le compartiment dendritique) ainsi que les mécanismes de régulation de l'interaction miR-134 : : Limk1 restent encore à explorer.

miR-127 et le proto-oncogène BCL6 –

miR-127 est localisé dans un îlot CpG méthylé dans la plupart des tissus (excepté dans le sperme) et son expression est fortement induite ($\times 49$) à la suite de traitements chimiques qui simultanément inhibent la méthylation de l'ADN et les activités histone désacétylases. Cet effet est décrit dans certaines lignées cancéreuses et tumeurs primaires mais pas dans des fibroblastes normaux. Il est à noter que l'expression de miR-136 et des autres miARN au proche voisinage de miR-127 n'est pas affectée, suggérant que dans ces lignées, miR-127 puisse être produit à partir de son propre promoteur. Enfin, miR-127 apparaît aussi capable de réguler négativement l'expression du proto-oncogène BCL6 (Saito *et al.*, 2006). Ces résultats laissent entrevoir la possibilité de

réactiver l'expression de miR-127 par « des drogues épigénétiques » dans des traitements anti-cancéreux.

La famille miR-376 et l'édition de l'ARN A vers I - Récemment, Kawahara *et al.*, (2007) ont montré que les pré-microARN apparentés à la famille miR-376 pouvaient être édités dans le cerveau par ADAR1 et ADAR2 au niveau de la position +4. Cette conversion A vers I, positionnée dans la région 5' fonctionnelle du miARN change sa spécificité de reconnaissance vis-à-vis de ces cibles. Ainsi le gène PRPS1, impliqué dans le métabolisme de l'uridine et la voie de synthèse de l'acide urique, est une cible des formes miR-376 éditées mais pas de celles non-éditées (Kawahara *et al.*, 2007). Comme on s'y attendait, les auteurs montrent que le niveau d'expression de la protéine PPS1 ainsi que le contenu d'acide urique dans le cortex de souris k.o. Adar2 -/- (absence de miR-376 édités) sont anormalement élevés par rapport à des souris sauvages (d'un facteur 2).

Conclusion :

Après avoir identifié et caractérisé l'organisation génique et l'expression des gènes des ARN C/D et microARN soumis à l'EGP, nombreuses sont les questions fondamentales qui persistent (Seitz *et al.*, 2003; Royo *et al.*, 2006). Par exemple, quelles sont les voies de signalisation qui gouvernent leur expression spécifique de tissus et/ou au cours du développement embryonnaire? Sont-ils impliqués dans des régulations épigénétiques? Quelle est la signification évolutive de ce mode d'expression mono-allélique et surtout quelles sont les voies physiologiques dans lesquels ils interviennent? Les travaux en cours au laboratoire visent à créer des modèles murins, notamment une invalidation ciblée des gènes des microARN (recombinaison homologue médiée via le système Cre-LoxP). Les gènes soumis à l'EGP jouent un rôle important au cours du **développement des mammifères placentaires** en régulant la croissance embryonnaire, les échanges fœtus-placenta et certains traits comportementaux chez l'adulte (Abu-Amero *et al.*, 2006; Smith *et al.*, 2006). Des travaux récents montrent qu'il existe un contrôle transcriptionnel des gènes soumis à l'EGP, notamment via le facteur de transcription Zac1 dont le gène est lui-même soumis à empreinte (Varrault *et al.*, 2006) et/ou vraisemblablement à travers des communications intra-chromosomiques entre loci (Zhao *et al.*, 2006). Il est par conséquent tentant de spéculer que les petits ARN décrits dans cet article, mais aussi ceux décrits à d'autres loci (Royo *et al.*, 2006), puissent moduler l'expression des gènes soumis à l'EGP (ou ceux impliqués dans les mêmes voies physiologiques). L'existence **d'un tel réseau post-transcriptionnel** via des petits ARN est largement spéculative mais la capacité des miARN à mettre sous silence Rtl1 illustre à merveille cette hypothèse.

Remerciements :

Je remercie particulièrement Patrice Vitali et Marie-Line Cavallé pour la relecture attentive du manuscrit et Guillaume Canal pour ses précieux conseils concernant la réalisation de la figure 1. Je m'excuse au près de certains collègues ne pas avoir mentionné leurs travaux en raison du manque de place. Notre programme de recherche est actuellement financé par deux contrats européens (STREP CallimiR et PWS dans le cadre FP6) ainsi que par l'EMBO *Young Investigator Programme* 2006.

Références

- Abu-Amero, S., Monk, D., Apostolidou, S., Stanier, P., and Moore, G. (2006). Imprinted genes and their role in human fetal growth. *Cytogenet Genome. Res.* 113, 262-270.
- Bachellerie, J.P., Cavaille, J., and Huttenhofer, A. (2002). The expanding snoRNA world. *Biochimie* 84, 775-790.
- Bagga, S., Bracht, J., Hunter, S., Massirer, K., Holtz, J., Eachus, R., and Pasquinelli, A.E. (2005). Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation. *Cell.* 122, 553-563.
- Bartel, D.P., and Chen, C.Z. (2004). Micromanagers of gene expression : the potentially widespread influence of metazoan microRNAs. *Nat. Rev. Genet.* 5, 396-400.
- Bernstein, B.E., Meissner, A., and Lander, E.S. (2007). The Mammalian epigenome. *Cell.* 128, 669-681.
- Bernstein, E., and Allis, C.D. (2005). RNA meets chromatin. *Genes. Dev.* 19, 1635-1655.
- Burger, J., Horn, D., Tonnies, H., Neitzel, H., and Reis, A. (2002). Familial interstitial 570 kbp deletion of the UBE3A gene region causing Angelman syndrome but not Prader-Willi syndrome. *Am. J. Med. Genet.* 111, 233-237.
- Burns, C.M., Chu, H., Rueter, S.M., Hutchinson, L.K., Canton, H., Sanders-Bush, E., and Emeson, R.B. (1997). Regulation of serotonin-2C receptor G-protein coupling by RNA editing. *Nature* 387, 303-308.
- Carninci, P., and Hayashizaki, Y. (2007). Noncoding RNA transcription beyond annotated genes. *Curr. Opin. Genet. Dev.*
- Cavaille, J., Buiting, K., Kiefmann, M., Lalonde, M., Brannan, C.I., Horsthemke, B., Bachellerie, J.P., Brosius, J., and Huttenhofer, A. (2000). Identification of brain-specific and imprinted small nucleolar RNA genes exhibiting an unusual genomic organization. *Proc. Natl. Acad. Sci. U S A* 97, 14311-14316.
- Cavaille, J., Nicoloso, M., and Bachellerie, J.P. (1996). Targeted ribose methylation of RNA in vivo directed by tailored antisense RNA guides. *Nature* 383, 732-735.
- Cavaille, J., Seitz, H., Paulsen, M., Ferguson-Smith, A.C., and Bachellerie, J.P. (2002). Identification of tandemly-repeated C/D snoRNA genes at the imprinted human 14q32 domain reminiscent of those at the Prader-Willi/Angelman syndrome region. *Hum. Mol. Genet.* 11, 1527-1538.
- Cavaille, J., Vitali, P., Basyuk, E., Huttenhofer, A., and Bachellerie, J.P. (2001). A novel brain-specific box C/D small nucleolar RNA processed from tandemly repeated introns of a noncoding RNA gene in rats. *J. Biol. Chem.* 276, 26374-26383.
- Chamberlain, S.J., and Brannan, C.I. (2001). The Prader-Willi syndrome imprinting center activates the paternally expressed murine Ube3a antisense transcript but represses paternal Ube3a. *Genomics* 73, 316-322.
- Charlier, C., Segers, K., Karim, L., Shay, T., Gyapay, G., Cockett, N., and Georges, M. (2001). The callipyge mutation enhances the expression of coregulated imprinted genes in cis without affecting their imprinting status. *Nat. Genet.* 27, 367-369.
- Davis, E., Caiment, F., Tordoier, X., Cavaille, J., Ferguson-Smith, A., Cockett, N., Georges, M., and Charlier, C. (2005). RNAi-mediated allelic trans-interaction at the imprinted Rtl1/Peg11 locus. *Curr. Biol.* 15, 743-749.
- Davis, E., Jensen, C.H., Schroder, H.D., Farnir, F., Shay-Hadfield, T., Kliem, A., Cockett, N., Georges, M., and Charlier, C. (2004). Ectopic expression of DLK1 protein in skeletal muscle of padumnal heterozygotes causes the callipyge phenotype. *Curr. Biol.* 14, 1858-1862.
- Ding, F., Prints, Y., Dhar, M.S., Johnson, D.K., Garnacho-Montero, C., Nicholls, R.D., and Francke, U. (2005). Lack of Pwcr1/MBII-85 snoRNA is critical for neonatal lethality in Prader-Willi syndrome mouse models. *Mamm Genome* 16, 424-431.
- Feil, R. (2006). Environmental and nutritional effects on the epigenetic regulation of genes. *Mutat. Res.* 600, 46-57.
- Freking, B.A., Murphy, S.K., Wylie, A.A., Rhodes, S.J., Keele, J.W., Leymaster, K.A., Jirtle, R.L., and Smith, T.P. (2002). Identification of the single base change causing the callipyge muscle hypertrophy phenotype, the only known example of polar overdominance in mammals. *Genome Res.* 12, 1496-1506.
- Gallagher, R.C., Pils, B., Albalwi, M., and Francke, U. (2002). Evidence for the role of PWCR1/HBII-85 C/D box small nucleolar RNAs in Prader-Willi syndrome. *Am. J. Hum. Genet.* 71, 669-678.
- Georges, M., Charlier, C., and Cockett, N. (2003). The callipyge locus : evidence for the trans interaction of reciprocally imprinted genes. *Trends Genet.* 19, 248-252.
- Kawahara, Y., Zinshteyn, B., Sethupathy, P., Iizasa, H., Hatzigeorgiou, A.G., and Nishikura, K. (2007). Redirection of silencing targets by adenosine-to-inosine editing of miRNAs. *Science* 315, 1137-1140.
- Kim, V.N. (2006). Small RNAs just got bigger : Piwi-interacting RNAs (piRNAs) in mammalian testes. *Genes. Dev.* 20, 1993-1997.
- Kishore, S., and Stamm, S. (2006). The snoRNA HBII-52 regulates alternative splicing of the serotonin receptor 2C. *Science* 311, 230-232.
- Kiss, T. (2002). Small nucleolar RNAs : an abundant group of noncoding RNAs with diverse cellular functions. *Cell.* 109, 145-148.

- Kiss-Laszlo, Z., Henry, Y., Bachelier, J.P., Caizergues-Ferrer, M., and Kiss, T. (1996). Site-specific ribose methylation of preribosomal RNA : a novel function for small nucleolar RNAs. *Cell*. 85, 1077-1088.
- Kloosterman, W.P., and Plasterk, R.H. (2006). The diverse functions of microRNAs in animal development and disease. *Dev. Cell*. 11, 441-450.
- Kouzarides, T. (2007). Chromatin modifications and their function. *Cell*. 128, 693-705.
- Landers, M., Bancescu, D.L., Le Meur, E., Rougeulle, C., Glatt-Deeley, H., Brannan, C., Muscatelli, F., and Lalande, M. (2004). Regulation of the large (approximately 1000 kb) imprinted murine Ube3a antisense transcript by alternative exons upstream of Snurf/Snrpn. *Nucleic Acids Res.* 32, 3480-3492.
- Le Meur, E., Watrin, F., Landers, M., Sturny, R., Lalande, M., and Muscatelli, F. (2005). Dynamic developmental regulation of the large non-coding RNA associated with the mouse 7C imprinted chromosomal region. *Dev. Biol.* 286, 587-600.
- Lewis, A., and Redrup, L. (2005). Genetic imprinting : conflict at the Callipyge locus. *Curr. Biol.* 15, R291-294.
- Lim, L.P., Lau, N.C., Garrett-Engle, P., Grimson, A., Schelter, J.M., Castle, J., Bartel, D.P., Linsley, P.S., and Johnson, J.M. (2005). Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* 433, 769-773.
- Lin, S.P., Youngson, N., Takada, S., Seitz, H., Reik, W., Paulsen, M., Cavaille, J., and Ferguson-Smith, A.C. (2003). Asymmetric regulation of imprinting on the maternal and paternal chromosomes at the Dlk1-Gtl2 imprinted cluster on mouse chromosome 12. *Nat. Genet.* 35, 97-102.
- Mancini-Dinardo, D., Steele, S.J., Levorse, J.M., Ingram, R.S., and Tilghman, S.M. (2006). Elongation of the Kcnq1ot1 transcript is required for genomic imprinting of neighboring genes. *Genes. Dev.* 20, 1268-1282.
- Mattick, J.S., and Makunin, I.V. (2006). Non-coding RNA. *Hum. Mol. Genet.* 15 *Spec No 1*, R17-29.
- Moore, T., and Haig, D. (1991). Genomic imprinting in mammalian development : a parental tug-of-war. *Trends Genet.* 7, 45-49.
- Nicholls, R.D., and Knepper, J.L. (2001). Genome organization, function, and imprinting in Prader-Willi and Angelman syndromes. *Annu. Rev. Genomics Hum. Genet.* 2, 153-175.
- Niswender, C.M., Copeland, S.C., Herrick-Davis, K., Emeson, R.B., and Sanders-Bush, E. (1999). RNA editing of the human serotonin 5-hydroxytryptamine 2C receptor silences constitutive activity. *J. Biol. Chem.* 274, 9472-9478.
- Niwa, R., and Slack, F.J. (2007). The evolution of animal microRNA function. *Curr. Opin. Genet. Dev.*
- Nottrott, S., Simard, M.J., and Richter, J.D. (2006). Human let-7a miRNA blocks protein production on actively translating polyribosomes. *Nat. Struct. Mol. Biol.* 13, 1108-1114.
- O'Neill, M.J. (2005). The influence of non-coding RNAs on allele-specific gene expression in mammals. *Hum. Mol. Genet.* 14 *Spec No 1*, R113-120.
- Petersen, C.P., Bordeleau, M.E., Pelletier, J., and Sharp, P.A. (2006). Short RNAs repress translation after initiation in mammalian cells. *Mol. Cell*. 21, 533-542.
- Pillai, R.S. (2005). MicroRNA function : multiple mechanisms for a tiny RNA ? *Rna* 11, 1753-1761.
- Pillai, R.S., Bhattacharyya, S.N., Artus, C.G., Zoller, T., Cougot, N., Basyuk, E., Bertrand, E., and Filipowicz, W. (2005). Inhibition of translational initiation by Let-7 MicroRNA in human cells. *Science* 309, 1573-1576.
- Pillai, R.S., Bhattacharyya, S.N., and Filipowicz, W. (2006). Repression of protein synthesis by miRNAs : how many mechanisms ? *Trends Cell. Biol.*
- Reik, W., and Walter, J. (2001). Genomic imprinting : parental influence on the genome. *Nat. Rev. Genet.* 2, 21-32.
- Rougeulle, C., Cardoso, C., Fontes, M., Colleaux, L., and Lalande, M. (1998). An imprinted antisense RNA overlaps UBE3A and a second maternally expressed transcript. *Nat. Genet.* 19, 15-16.
- Royo, H., Bortolin, M.L., Seitz, H., and Cavaille, J. (2006). Small non-coding RNAs and genomic imprinting. *Cytogenet Genome Res.* 113, 99-108.
- Runte, M., Huttenhofer, A., Gross, S., Kiefmann, M., Horsthemke, B., and Buiting, K. (2001). The IC-SNURF-SNRPN transcript serves as a host for multiple small nucleolar RNA species and as an antisense RNA for UBE3A. *Hum. Mol. Genet.* 10, 2687-2700.
- Runte, M., Varon, R., Horn, D., Horsthemke, B., and Buiting, K. (2005). Exclusion of the C/D box snoRNA gene cluster HBII-52 from a major role in Prader-Willi syndrome. *Hum. Genet.* 116, 228-230.
- Saito, Y., Liang, G., Egger, G., Friedman, J.M., Chuang, J.C., Coetzee, G.A., and Jones, P.A. (2006). Specific activation of microRNA-127 with downregulation of the proto-oncogene BCL6 by chromatin-modifying drugs in human cancer cells. *Cancer. Cell*. 9, 435-443.
- Schratt, G.M., Tuebing, F., Nigh, E.A., Kane, C.G., Sabatini, M.E., Kiebler, M., and Greenberg, M.E. (2006). A brain-specific microRNA regulates dendritic spine development. *Nature* 439, 283-289.
- Schule, B., Albalwi, M., Northrop, E., Francis, D.I., Rowell, M., Slater, H., McKinlay Gardner, R.J., and Francke, U. (2005). Molecular breakpoint cloning and gene expression studies of a novel translocation t(4;15)(q27;q11.2) associated with Prader-Willi syndrome. *BMC Med. Genet.* 6, 18.
- Seidl, C.I., Stricker, S.H., and Barlow, D.P. (2006). The imprinted Air ncRNA is an atypical RNAPII transcript that evades splicing and escapes nuclear export. *Embo. J.* 25, 3565-3575.
- Seitz, H., Royo, H., Bortolin, M.L., Lin, S.P., Ferguson-Smith, A.C., and Cavaille, J. (2004a). A large imprinted microRNA gene cluster at the mouse Dlk1-Gtl2 domain. *Genome. Res.* 14, 1741-1748.
- Seitz, H., Royo, H., Lin, S.P., Youngson, N., Ferguson-Smith, A.C., and Cavaille, J. (2004b).

- Imprinted small RNA genes. *Biol. Chem.* 385, 905-911.
- Seitz, H., Youngson, N., Lin, S.P., Dalbert, S., Paulsen, M., Bachelier, J.P., Ferguson-Smith, A.C., and Cavaille, J. (2003). Imprinted microRNA genes transcribed antisense to a reciprocally imprinted retrotransposon-like gene. *Nat. Genet.* 34, 261-262.
- Sleutels, F., Zwart, R., and Barlow, D.P. (2002). The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature* 415, 810-813.
- Smit, M., Segers, K., Carrascosa, L.G., Shay, T., Baraldi, F., Gyapay, G., Snowden, G., Georges, M., Cockett, N., and Charlier, C. (2003). Mosaicism of Solid Gold supports the causality of a noncoding A-to-G transition in the determinism of the callipyge phenotype. *Genetics* 163, 453-456.
- Smith, F.M., Garfield, A.S., and Ward, A. (2006). Regulation of growth and metabolism by imprinted genes. *Cytogenet. Genome Res.* 113, 279-291.
- Surani, M.A., Hayashi, K., and Hajkova, P. (2007). Genetic and epigenetic regulators of pluripotency. *Cell* 128, 747-762.
- Tai, H.C., and Schuman, E.M. (2006). MicroRNA : microRNAs reach out into dendrites. *Curr. Biol.* 16, R121-123.
- Varrault, A., Gueydan, C., Delalbre, A., Bellmann, A., Houssami, S., Akin, C., Severac, D., Chotard, L., Kahli, M., Le Digarcher, A., Pavlidis, P., and Journot, L. (2006). Zac1 regulates an imprinted gene network critically involved in the control of embryonic growth. *Dev. Cell* 11, 711-722.
- Vitali, P., Basyuk, E., Le Meur, E., Bertrand, E., Muscatelli, F., Cavaille, J., and Huttenhofer, A. (2005). ADAR2-mediated editing of RNA substrates in the nucleolus is inhibited by C/D small nucleolar RNAs. *J. Cell. Biol.* 169, 745-753.
- Wang, Q., O'Brien, P.J., Chen, C.X., Cho, D.S., Murray, J.M., and Nishikura, K. (2000). Altered G protein-coupling functions of RNA editing isoform and splicing variant serotonin_{2C} receptors. *J. Neurochem.* 74, 1290-1300.
- Wirth, J., Back, E., Huttenhofer, A., Nothwang, H.G., Lich, C., Gross, S., Menzel, C., Schinzel, A., Kioschis, P., Tommerup, N., Ropers, H.H., Horsthemke, B., and Buiting, K. (2001). A translocation breakpoint cluster disrupts the newly defined 3' end of the SNURF-SNRPN transcription unit on chromosome 15. *Hum. Mol. Genet.* 10, 201-210.
- Yang, P.K., and Kuroda, M.I. (2007). Noncoding RNAs and Intranuclear Positioning in Monoallelic Gene Expression. *Cell* 128, 777-786.
- Yekta, S., Shih, I.H., and Bartel, D.P. (2004). MicroRNA-directed cleavage of HOXB8 mRNA. *Science* 304, 594-596.
- Youngson, N.A., Kocialkowski, S., Peel, N., and Ferguson-Smith, A.C. (2005). A small family of sushiclass retrotransposon-derived genes in mammals and their relation to genomic imprinting. *J. Mol. Evol.* 61, 481-490.
- Zamore, P.D., and Haley, B. (2005). Ribo-gnome : the big world of small RNAs. *Science* 309, 1519-1524.
- Zhao, Z., Tavoosidana, G., Sjolinder, M., Gondor, A., Mariano, P., Wang, S., Kanduri, C., Lezcano, M., Sandhu, K.S., Singh, U., Pant, V., Tiwari, V., Kurukuti, S., and Ohlsson, R. (2006). Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and inter-chromosomal interactions. *Nat. Genet.* 38, 1341-1347.