

Identification et caractérisation de longs ARN non codants exprimés dans les neurones dopaminergiques et potentiellement impliqués dans la maladie de Parkinson

Claire Sauty-Colace

▶ **To cite this version:**

Claire Sauty-Colace. Identification et caractérisation de longs ARN non codants exprimés dans les neurones dopaminergiques et potentiellement impliqués dans la maladie de Parkinson. Biologie moléculaire. 2018. hal-01946848

HAL Id: hal-01946848
https://ephe.hal.science/hal-01946848

Submitted on 17 Dec 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE
ECOLE PRATIQUE DES HAUTES ETUDES
Sciences de la Vie et de la Terre

MEMOIRE présenté par

Claire COLACE-SAUTY
Pour l'obtention du Diplôme de l'Ecole Pratique des Hautes Etudes

Identification et caractérisation de longs ARN non codants exprimés dans les neurones dopaminergiques et potentiellement impliqués dans la maladie de Parkinson

Présenté et soutenu devant le Jury composé de :
Pr. Bernard MIGNOTTE - Président
Dr. Hélène CHEVAL – Tuteur scientifique
Dr. Khalid EL HACHIMI – Tuteur pédagogique
Dr. Alexis BEMELMANS – Rapporteur
Dr. Jeannette NARDELLI – Examinateur

Mémoire préparé sous la direction de :
Dr. CHEVAL Hélène
Directeur : Dr. Philippe RAVASSARD
Laboratoire d'accueil :
Equipe Biotechnologies et Biothérapies,
ICM, UPMC/Inserm UMR S975/CNRS UMR 7225
GH Pitié Salpêtrière,
47 bd de l'Hôpital, 75013 PARIS

Et de :
Dr. Khalid H. El Hachimi
Directeur : Dr. Stevanin Giovanni
Laboratoire EPHE :
Laboratoire de Neurogénétique EPHE
ICM, UPMC/Inserm UMR S975/CNRS UMR 7225
GH Pitié Salpêtrière,
47 bd de l'Hôpital, 75013 PARIS
EPHE (Sciences de la Vie et de la Terre)
Groupe de Recherche et d'Enseignement
Thématiques de l'EPHE : Trans-Neuro

Table des matières

Table des matières 2
Liste des illustrations 3
Liste des tableaux 3
Liste des principales abréviations 4
Introduction 5
A La maladie de Parkinson 6
1 Epidémiologie 6
2 Diagnostic clinique de la MP 6
3 Traitements 11
4 Etiologie 14
B Les longs ARN non codants 20
1 Définition d'un long ARN non-codant 21
2 Classification des lncARN 23
3 Conservation 25
4 Localisation cellulaire 26
5 Mécanismes d'action 27
6 Implication des lncARN dans les pathologies humaines 31
Références bibliographiques 34

Liste des illustrations

Figure 1 : Neuropathologie de la maladie de Parkinson 7
Figure 2 : Représentation schématique du circuit des ganglions de la base..... 8
Figure 3 : Symptômes cliniques et progression de la MP 9
Figure 4 : Marquage immunohistochimique d'un corps de Lewy dans un neurone DA de la SN*pc*..... 10
Figure 5 : Propagation des corps de Lewy au cours du temps dans la maladie de Parkinson..... 10
Figure 6 : Schéma général de la biosynthèse de l'ensemble des catécholamines 11
Figure 7 : Représentation schématique du métabolisme du MPTP..... 14
Figure 8 : Représentation schématique du mécanisme d'agrégation de l'α-synucléine..... 17
Figure 9 : Découverte de 26 loci à risques associés à la maladie de Parkinson 18
Figure 10 : Répartition des transcrits du génome humain..... 20
Figure 11 : Organisation chromosomique des lncARN par rapport au gène le plus proche 24
Figure 12 : Exemple de séquence conservée pour le LincARN OIP5-AS1 ou CYRANO 26
Figure 13 : Mode d'action en cis ou trans des lncARN..... 28
Figure 14 : Mécanismes de régulation des lncARN..... 28
Figure 15 : Représentation schématique du mécanisme d'action du lncARN BACE1-AS..... 32

Liste des tableaux

Tableau 1 : Gènes impliqués dans les formes familiales de la maladie de Parkinson 16

Liste des principales abréviations

6-OHDA	6-hydroxydopamine
aa	Acide aminé
ADN	Acide désoxyribonucléique
ADNc	ADN complémentaire
ADNg	ADN génomique
ARN	Acide ribonucléique
ARNm	ARN messager
ARNr	ARN ribosomal
ARNt	ARN de transfert
ATAC-Seq	Séquençage des zones ouvertes de la chromatine
BMP	Bone Morphogenetic protein
ceARN	Competing endogenous RNA
DA	Dopaminergique
DAT	Transporteur de la dopamine
DBH	Dopamine β-hydroxylase
DDC	Dopa décarboxylase
DRD2	Récepteur de la Dopamine D2
GDNF	Glial Derived Neurotrophic Factor
GWAS	Genome Wide Association Study
hiPSC	Cellules pluripotentes induites humaines
L-DOPA	Lévodopa
LincARN	LncARN intergénique
lncARN	Longs ARN non codants
lncARN-AS	lncARN anti-sens
MA	Maladie d'Alzheimer
MALAT1	Metastasis-Associated Lung Adenocarcinoma Transcript 1
MAPT	Microtubule associated protein tau
MH	Maladie de Huntington
miRNA	microARN
MP	Maladie de Parkinson
MPTP	1-méthyl-4-phényl-1,2,3,6-tétrahydropyridine
NA	Noradrénergiques
ncARN	ARN non codants
NEAT1	Nuclear Paraspeckle Assembly Transcript 1
PINK1	PTEN Induced Putative Kinase 1
RNA-Seq	Séquençage de l'ARN
SNCA	gene de l'alpha synucléine
SNHG1	Small nucleolar RNA host gene 1
SN*pc*	Substance noire *pars compacta*
TH	Tyrosine hydroxylase
TSS	Site de démarrage de la transcription
VMAT2	vesicular monoamine transporter

Introduction

A La maladie de Parkinson

Le médecin britannique James Parkinson publie pour la première fois en 1817 une description des symptômes de « la paralysie agitante » qui sera redéfinie comme la maladie de Parkinson (MP) 55 années plus tard par le célèbre neurobiologiste français Jean-Martin Charcot (Parkinson, 2002). La maladie de Parkinson est une affection neurodégénérative qui se caractérise essentiellement par une perte progressive des neurones dopaminergiques (DA) de la substance noire *pars compacta* (SN*pc*). Compte tenu du vieillissement de la population et du nombre grandissant de patients parkinsoniens, la maladie de Parkinson constitue un véritable problème de santé publique et il est important d'essayer de mieux comprendre l'étiologie de la maladie.

1 Epidémiologie

Le recensement des patients souffrant de la MP n'est pas aisé. Les premières études épidémiologiques ont été réalisées sous la forme de grandes enquêtes systématiques (porte à porte) dans une large population afin d'évaluer le pourcentage de la population générale atteinte par la MP. L'avantage de ces approches était de pouvoir détecter des patients non diagnostiqués. Toutefois, ces études couteuses étaient difficiles à mettre en œuvre et ne concernaient qu'une population restreinte (Elbaz et al., 2016). De plus, il a été montré que le choix des critères de diagnostic excluant d'autres causes de parkinsonisme, comme par exemple la réponse à un traitement médicamenteux, peut influencer l'estimation de personnes atteintes de la MP (de Rijk et al., 1997). Les certificats de décès dans lesquels la maladie n'est pas toujours renseignée, ne constituent pas un bon outil pour l'évaluation du nombre de patients Parkinsoniens (Beyer et al., 2001). En revanche, les bases de données de demandes de remboursement de médicaments sont une source intéressante pour l'identification des patients atteints de formes spécifiques de la MP. Cependant, les médicaments antiparkinsoniens n'étant pas exclusivement spécifiques à la maladie de Parkinson, la validation des données reste une étape importante. Une étude en France montre que l'identification des patients parkinsoniens utilisant des médicaments antiparkinsoniens est une approche valable, notamment lorsque plusieurs médicaments et dosages sont pris en compte (Moisan et al., 2011).

La maladie de Parkinson reste relativement rare bien qu'elle soit la seconde maladie neurodégénérative la plus fréquente après la maladie d'Alzheimer (Dorsey et al., 2007). Dans le monde, elle touche entre 10 à 30 individus sur 10 000 personnes et chaque année, 1 à 5 nouveaux cas sont recensés pour une population de 10 000 personnes (Elbaz et al., 2016; von Campenhausen et al., 2005). Compte tenu du vieillissement de la population, ces chiffres devraient doubler d'ici 2030 (Dorsey et al., 2007). La fréquence de la MP augmente progressivement et fortement avec l'âge, elle est rare avant l'âge de 50 ans et atteint 1 % de la population après 70 ans (Pringsheim et al., 2014). Elle touche environ 1,5 fois plus les hommes que les femmes (Wooten et al., 2004). En France, la prévalence est de 25 patients pour 10 000 et l'incidence est de 3,9 nouveaux cas pour 10 000 pour l'année 2015, à partir des données de remboursement des médicaments antiparkinsoniens du Système national d'assurance maladie (Moisan et al., 2018).

2 Diagnostic clinique de la MP

La MP est une maladie neurodégénérative chronique à progression lente. A l'heure actuelle, son diagnostic est essentiellement clinique, bien que récemment une analyse morphométrique d'IRM (Imagerie à Résonnance Magnétique) de patients parkinsoniens a permis d'identifier une signature

cérébrale liée à l'atrophie du cerveau et représentant la gravité de la MP et de prédire la progression de la maladie motrice et cognitive chez de nouveaux patients (Zeighami et al., 2017). Ainsi, il existe un panel de symptômes moteurs et non-moteurs très invalidants et qui varient beaucoup en fonction des patients, rendant le diagnostic clinique difficile. En 2015, les étapes de la pathologie sont définies en trois catégories par la société internationale contre la maladie de Parkinson et les troubles du mouvement (Kalia and Lang, 2016):

- la MP préclinique, lorsque la neurodégénérescence a probablement débuté sans que le patient ne présente encore de symptômes,
- la MP prodromique, lorsque les signes cliniques apparaissent mais sont insuffisants pour diagnostiquer une MP clinique (Berg et al., 2015),
- la MP clinique, lorsque tous les critères de diagnostiques sont réunis.

Étant donné l'absence de traitement neuroprotecteur de la MP, les deux premières catégories ont été élaborées uniquement à des fins de recherche.

2-1 Symptômes moteurs et régions cérébrales associées

Le diagnostic clinique de la MP repose sur une triade de symptômes moteurs classiquement observés (Postuma et al., 2015):

- une akinésie (lenteur dans l'initiation et la coordination de mouvements), entrainant des difficultés notamment pour la marche ou l'écriture,
- une hypertonie (rigidité musculaire excessive) provoquant des douleurs musculaires et une altération de la posture,
- un tremblement au repos touchant principalement les membres supérieurs et souvent asymétrique c'est-à-dire d'un seul côté du corps.

Certains troubles peuvent apparaitre plus tardivement lorsque la maladie est à un stade plus avancé comme le blocage de la marche ou « freezing », des problèmes d'équilibre entrainant des chutes et des troubles de la déglutition. L'apparition de ces troubles moteurs est très hétérogène selon les patients (Kalia and Lang, 2015).

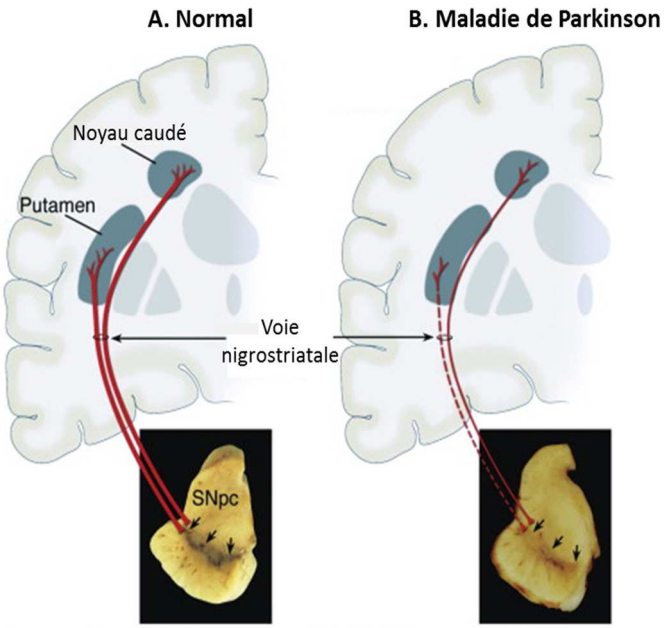


Figure 1 : Neuropathologie de la maladie de Parkinson (adapté de Dauer, 2003).
(A) Représentation de la voie normale nigrostriatale (lignes rouges épaisses). Elle est composée de neurones DA dont les corps cellulaires sont localisés dans la substance noire pars compacta (SN*pc*, flèches noires). Ces neurones projettent jusqu'au striatum (putamen et noyau caudé). La photo montre la pigmentation normale de la SN*pc*, produite par la mélanine dans les neurones DA. (B) Représentation schématique de la dégénerescence de la voie nigrostriatale dans la MP. Il y a une perte marquée de neurones DA qui projettent vers le putamen (ligne rouge pointillée) et une perte beaucoup plus modeste de ceux qui projettent vers le noyau caudé (fine ligne rouge continue). La photographie montre une dépigmentation de la SN*pc* due à la perte marquée de neurones DA (flèches noires).

En 1919, Konstantin Nikolaevich Tretiakoff montra une dépigmentation de la substance noire due à la dégénérescence des neurones DA mélanisés dans la MP et proposa un lien entre la perte de cellules nerveuses et les symptômes moteurs (rigidité musculaire et tremblements) (Goedert et al., 2013). Aujourd'hui, le fait que ces neurones dans la partie ventrale de la SN*pc* soient sévèrement affectés par la MP est maintenant bien établi (Figure 1). Cette région, associée aux ganglions de la base, est impliquée dans le contrôle des mouvements volontaires, en particulier dans la phase d'initiation des mouvements et dans leur coordination. Les neurones DA projettent principalement vers le putamen et le noyau caudé, les régions du striatum les plus gravement appauvries en dopamine dans la MP et responsables des principaux symptômes moteurs de la maladie (Figure 2) (Maiti et al., 2017).

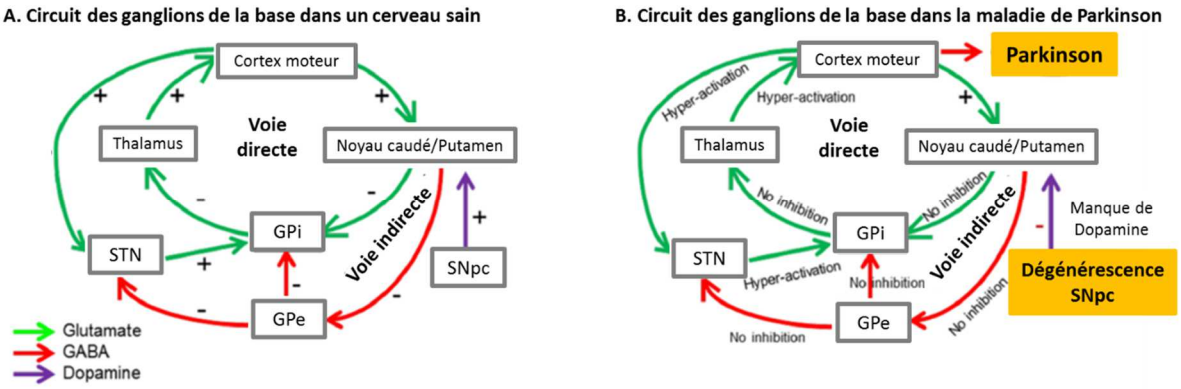


Figure 2 : Représentation schématique du circuit des ganglions de la base chez un individu sain (A) ou un patient atteint de la maladie de Parkinson (B) (adapté de Maiti, 2017).

Les signes (+) représentent une activation et les signes (-) représentent une inhibition. La perte des neurones DA qui affèrent vers le putamen et le noyau caudé altère la voie nigro-striatale. Il n'y a alors plus d'effet inhibiteurs (No inhibition) sur le *globus pallidus* et le thalamus, ce qui entraine une hyper-activation du cortex moteur, provoquant une altération de la coordination et du contrôle des mouvements chez les patients atteints de la MP. GPi : *Globus pallidus* interne, Gpe : *Globus pallidus* externe, SN*pc* : Substance noire *pars compacta*, STN : noyau sous-thalamique.

2-2 Symptômes non moteurs et régions cérébrales associées

Les troubles non moteurs peuvent apparaitre 20 ans avant l'apparition des premiers symptômes moteurs, entrainant ainsi une diminution de la qualité de vie des patients (Kalia and Lang, 2015). Dans cette phase prodromique, les symptômes suivants peuvent se déclarer (Figure 3) (Kalia and Lang, 2015):

- des troubles du sommeil comme des insomnies, des somnolences diurnes (en journée) ou des troubles du comportement pendant la phase de sommeil paradoxal (Rapid eye movement sleep Behavior Disorder =RBD),
- des troubles digestifs liés à la perturbation du système autonome comme l'hypersalivation ou la constipation,
- des troubles olfactifs,
- des troubles psychologiques comme la dépression ou l'altération des fonctions exécutives (attention, planification, anticipation, concentration).

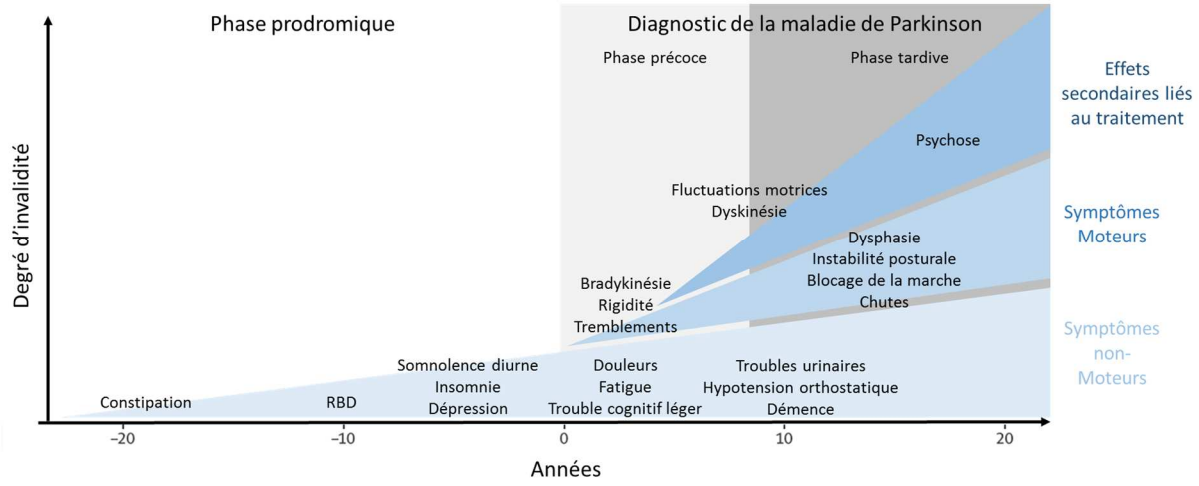


Figure 3 : Symptômes cliniques et progression de la MP (adapté de Kalia & Lang, 2015).

La MP ne peut plus être considérée comme une maladie mono-systémique qui affecte principalement le système DA nigrostriatal puisqu'elle affecte de nombreuses régions différentes du système nerveux. Bien qu'il soit difficile de démontrer pour chaque neurotransmetteur un rôle spécifique dans la physiopathologie de la maladie, certains systèmes ont pu être associés à la MP. Avec l'arrivée des premiers symptômes moteurs, peuvent apparaitre des douleurs musculaires et une fatigue excessive. Cette dernière peut être associée à un dysfonctionnement des neurones sérotoninergiques des noyaux du raphé, probablement dû à un métabolisme anormal du fer (Zuo et al., 2016). D'autres troubles d'ordre cognitif apparaissent plus tardivement, tels que l'anxiété, l'apathie, les troubles de l'attention et de la vigilance, et même la démence, et reflètent des lésions des systèmes noradrénergiques du *locus coeruleus* (Delaville et al., 2011). En plus des pertubations motrices, des troubles du sommeil de la MP ont été associés à une perte des les neurones cholinergiques du noyau pédonculopontin et du noyau basal de Meynert (Pepeu and Grazia Giovannini, 2017) et à une diminution des cellules à hypocrétine de l'hypothalamus (Thannickal et al., 2007).

Les symptômes observés sont très variables d'une personne à l'autre et peuvent varier selon divers facteurs comme le stress, les émotions et le sommeil.

2-3 Anatomie pathologique

Il y a plus d'un siècle, Frizt Jakob Heinrich Lewy étudia des cerveaux de patients parkinsoniens *post mortem* à travers la microscopie optique et découvrit la présence d'inclusions protéiques caractéristiques qui seront décrites par la suite dans la substance noire et nommées « corps de Lewy » par Konstantin Nikolaevich Tretiakoff (Goedert et al., 2013). Aujourd'hui encore, la principale caractéristique pathologique de la maladie de Parkinson est la formation d'inclusions cytoplasmiques appelées « corps et neurites de Lewy » dans les neurones survivants (Figure 4).

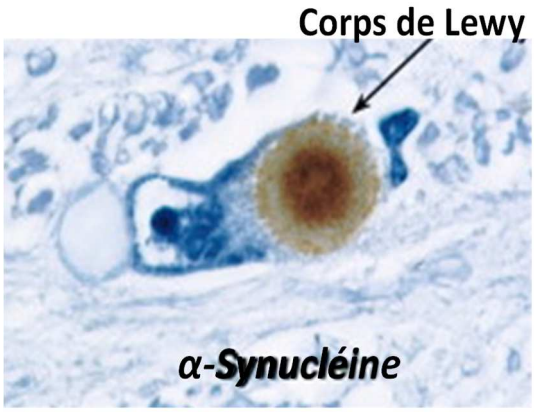


Figure 4 : Marquage immunohistochimique d'un corps de Lewy dans un neurone DA de la SN*pc* (adapté de Dauer, 2003).

Les corps de Lewy sont des agglomérats anormaux de protéines insolubles, intra-cytoplasmiques, de formes sphériques. La protéine majoritaire constituant les corps de Lewy est l'α-synucléine codée par le gène *SNCA* (voir le paragraphe Facteurs génétiques). Ces inclusions ne sont pas spécifiques à la MP car elles sont présentes dans d'autres pathologies neurodégénératives comme la maladie d'Alzheimer et la démence à corps de Lewy. La prévention de la formation de ces inclusions pathologiques est un objectif majeur pour les années à venir.

De plus, des études comparatives de cerveaux humains normaux et malades ont montré que les inclusions d'α-synucléine émergent dans un ordre prévisible dans différentes parties du cerveau, ce qui permet de distinguer six stades de dépôt d'α-synucléine dans la MP (Braak et al., 2003) (Figure 5) :

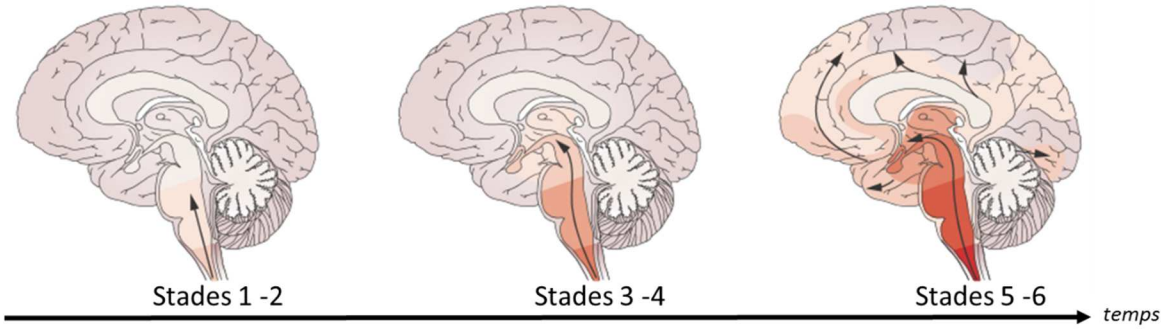


Figure 5 : Propagation des corps de Lewy au cours du temps dans la maladie de Parkinson (adapté de Goedert, 2013).

Stade 1 : apparition des premières inclusions d'α-synucléine dans le bulbe olfactif et / ou le noyau moteur dorsal du nerf vagal.
Stade 2 : la pathologie de Lewy se développe dans le bulbe rachidien et le tegmentum pontique et plus précisément dans les noyaux du raphé et du locus coeruleus.
Ces deux premiers stades correspondent à la phase prodromique d'apparition des symptômes non moteurs caractéristiques de la MP entrainant peu de perturbations de la vie du patient.
Stade 3 : la pathologie a atteint l'amygdale et la substance noire concordant avec l'apparition des symptômes moteurs de la maladie de Parkinson (bradykinésie, avec au moins une des trois caractéristiques de rigidité, de tremblement de repos ou de troubles de la marche), le patient est autonome dans sa vie quotidienne.
Stade 4 : la pathologie s'aggrave et les inclusions d'α-synucléine atteignent le cortex temporal. La marche devient difficile et l'autonomie est restreinte.
Stades 5 et 6 : les corps et les neurites de Lewy apparaissent dans le néocortex, ce qui explique de nombreux problèmes cognitifs associés à la maladie de Parkinson avancée. De plus, le patient n'est plus autonome et ne peut plus marcher.

L'accumulation de l'α-synucléine a aussi été décrite dans le système nerveux entérique plusieurs années avant l'apparition des premiers symptômes moteurs de la maladie de Parkinson (Shannon et al., 2012). Cependant, aucune analyse ne démontre le phénomène dynamique du transfert de l'α-synucléine d'une région à une autre (Lionnet et al., 2018). A ce jour, l'origine de la MP reste donc encore à élucider.

3 Traitements

La perte de neurones DA dans la SN*pc* conduit à une déplétion en dopamine dans le striatum et constitue la principale cause de l'apparition des symptômes moteurs de la MP. A ce jour, les traitements pharmacologiques existants sont essentiellement palliatifs et suivent deux axes : l'apport de dopamine et l'inhibition de sa dégradation.

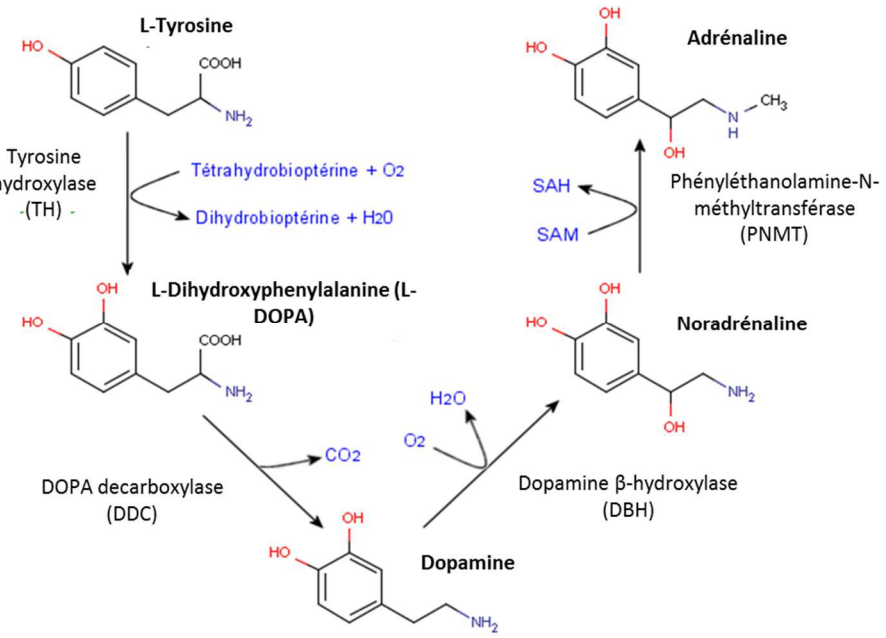


Figure 6 : Schéma général de la biosynthèse de l'ensemble des catécholamines.
Dans les neurones DA, les enzymes DBH et PNMT ne sont pas exprimées, la synthèse s'arrête à la dopamine.

Le traitement à la L-DOPA (ou Lévodopa), précurseur direct de la dopamine (Figure 6) a une efficacité quasi instantanée sur les symptômes moteurs (LeWitt and Fahn, 2016). Elle est administrée à l'aide d'une pompe placée dans l'intestin. Comme la L-DOPA est administrée en systémique, il est important d'éviter sa transformation en dopamine par la DDC (Dopa décarboxylase) ou sa dégradation avant d'atteindre le système nerveux central. Ainsi, la L-DOPA est associée à des inhibiteurs enzymatiques spécifiques (Fox et al., 2011):

- la Benserazide ou la Carbidopa inhibent la DDC périphérique. Comme ces molécules ne passent pas la barrière hématoencéphalique elles sont inactives sur la DDC centrale et n'interfèrent donc pas avec la production de dopamine dans le cerveau.
- l'Entacapone, l'Opicapone et la Tolcapone sont des inhibiteurs de catéchol-O-méthyltransférase (COMT) empêchant la méthylation de la L-DOPA la rendant ainsi plus disponible et prolongeant les effets de ces médicaments.
- la Rasagiline, la Selegiline et la Safinamide inhibent la dégradation de la dopamine en acide dihydroxyphénilacétique (DOPAC) et en peroxyde d'hydrogène (H_2O_2) par la monoamine oxydase de type B (MAOB) prolongeant ainsi l'effet de la dopamine dans la fente synaptique.

Cependant, l'administration prolongée de fortes doses de L-DOPA induit des fluctuations motrices et des dyskinésies (mouvements involontaires anormaux) chez la plupart des patients après plus de 5 ans de traitement (Cenci, 2014). De plus, la réponse à ce traitement évolue avec la progression de la maladie jusqu'à devenir inefficace sur les symptômes moteurs sans causer de dyskinésies. Utilisé en première intention chez les patients âgés de plus de 70 ans, le traitement à la L-DOPA peut aussi provoquer, à forte dose sur de plus jeunes patients, d'autres effets secondaires comme des nausées, des troubles cardiaques ou des troubles psychiatriques (Cenci, 2014). Dans ce cas et dans les stades précoces de la maladie, la stratégie thérapeutique s'oriente vers le choix d'agonistes dopaminergiques en association ou en substitution de la L-DOPA (Fox et al., 2011). Ce sont des médicaments capables de traverser la barrière hémato-encéphalique pour mimer directement l'action de la dopamine sur ses récepteurs postsynaptiques spécifiques :

- le Pramipexole, le Ropinirole et le Piribédil stimulent spécifiquement les récepteurs dopaminergiques de type D2 (D2R) de la voie indirecte .
- l'Apomorphine et la Rotigotine ciblent les récepteurs de type D2 mais aussi de type D1 (D1R) de la voie nigro-striée directe

Leur efficacité est moindre sur les symptômes moteurs de la MP, néanmoins ils permettent de repousser l'utilisation de L-DOPA chez les jeunes parkinsoniens et ainsi les effets secondaires associés. Cependant ces agonistes peuvent provoquer également des troubles psychologiques comme des hallucinations, ou des troubles de l'humeur. Les doses médicamenteuses sont augmentées proportionnellement à la progression de la maladie et associés à la L-DOPA lorsque leur efficacité diminue.

Malgré l'effet remarquable du traitement dopaminergique sur les symptômes de la maladie de Parkinson, l'étude d'autres systèmes pharmacologiques est nécessaire pour éviter les effets secondaires associés ou encore cibler d'autres systèmes neuronaux impliqués dans la MP. Actuellement, le seul traitement pharmacologique non dopaminergique efficace et disponible contre la dyskinésie induite par la L-DOPA est l'amantadine. Cette molécule empêcherait la recapture de la dopamine et provoquerait la libération de dopamine au niveau du striatum afin de corriger le déficit dopaminergique (Vanle et al., 2018). Mais de nombreux symptômes non moteurs peuvent être aggravés ou précipités par ce traitement. Les inhibiteurs de la cholinestérase peuvent avoir des effets bénéfiques sur les troubles cognitifs des patients parkinsoniens comme la récupération de fonctions cognitives, l'amélioration des troubles du comportement ou encore la réduction des hallucinations (Connolly and Fox, 2014).

Devant la diversité des patients souffrant de la MP, les technologies actuelles permettront d'identifier des groupes de patients homogènes d'un point de vue moléculaire pour pouvoir les traiter spécifiquement afin d'obtenir une meilleure réponse aux traitements.

Outre l'approche pharmacologique, il existe aussi une approche chirurgicale dans le traitement de la MP : la stimulation cérébrale profonde. Il s'agit d'une méthode invasive consistant à implanter dans le noyau sous-thalamique ou le *globus pallidus* des électrodes connectées à un boîtier placé sous la peau et délivrant un courant électrique de faible intensité. Ce traitement est proposé lorsque la MP est avancée à des patients traités à la L-DOPA mais présentant à long terme des dyskinésies et des fluctuations motrices (Bronstein et al., 2011). En conséquence, ces complications motrices peuvent diminuer de 70%, améliorant nettement le quotidien des patients (Deuschl and Agid, 2013). La neurostimulation semble efficace les 5 premières années puis, à plus long terme, les effets semblent

comparables à ceux observés chez les patients traités par médicaments (Gronchi-Perrin et al., 2009). Cependant, la chirurgie comporte des risques non négligeables qui sont les hémorragies intracrâniennes, les infections ou le mauvais positionnement de l'appareillage. De plus, les séquelles psychiatriques comme l'apathie ou les troubles de l'humeur ne sont pas rares et liées au déséquilibre dopaminergique dû aux changements de médication et aux effets induits par la stimulation (Volkmann et al., 2010). Des dispositifs de neurostimulation automatisés sont maintenant disponibles et permettent d'homogénéiser les résultats entre les centres de chirurgie comme par exemple l'utilisation de systèmes de neurostimulation en boucles fermées dont les paramètres s'auto-ajustent en fonction des marqueurs physiologiques (Kuhn and Volkmann, 2017).

En parallèle de ces traitements médicamenteux et chirurgicaux, la kinésithérapie est efficace pour le maintien et l'amélioration de la marche et de l'équilibre, et la rééducation orthophonique est utile pour corriger la déglutition, et les problèmes d'élocution.

Néanmoins, les traitements actuels n'empêchent pas la perte continue des différents types de neurones impliqués dans la MP. Aujourd'hui, le défi de la recherche est de trouver un traitement neuro-protecteur ou neuro-restaurateur pour stopper l'état de la MP ou rétablir les systèmes neuronaux endommagés. Une approche immunologique *via* un vaccin dirigé contre l'α-synucléine est actuellement en développement clinique. En effet, dans un modèle murin transgénique de la maladie de Parkinson, l'immunisation active avec un nouveau vaccin contenant des peptides courts homologues de l'α-synucléine conjuguée à un adjuvant (AFFITOPE, AFFiRiS) a induit la formation d'anticorps spécifiquement dirigés contre l'extrémité carboxy de l'α-synucléine humaine, permettant d'éliminer les agrégats de l'α-synucléine et de réduire la neuropathologie (Mandler et al., 2014). Un essai clinique de phase I / II chez 28 patients parkinsonniens a également montré la formation d'anticorps contre l'α-synucléine sans danger, et l'efficacité immunologique des vaccinations de rappel jusqu'à 3 ans après la vaccination initiale (McGuire Kuhl et al.,2016). Une autre approche de thérapie génique a vu le jour, basée sur le fait que l'apport de facteurs de la familles des GDNF (Glial Derived Neurotrophic Factor) comme la neurturine protègerait les neurones DA de la mort et favoriserait la neurorestauration des axones endommagés dans des modèles animaux (Brizard et al., 2006; Kordower and Bjorklund, 2013). Un premier essai clinique visant à restaurer le réseau dopaminergique lésé dans la MP en induisant l'expression de la neurturine par l'intermédiaire d'un vecteur AAV (Adeno-associated virus) n'a pas été validé car peu de cellules exprimaient le facteur neurotrophique, limitant ainsi son effet neurorestaurateur sur les neurones DA (Bartus and Johnson, 2017a, b). Un nouvel essai est en cours pour tester les effets de l'expression du GDNF par les AAV sur des patients en phase précoce.

Actuellement, le développement de nouveaux traitements cellulaires de la MP visant à ralentir ou inverser la progression de la maladie est en plein essor (Man et al., 2018). Des protocoles ont été développés pour générer des neurones DA en différenciant des cellules souches pluripotentes humaines provenant d'embryons (hESC) ou de cellules somatiques adultes reprogrammées (hiPSC) (Kirkeby et al., 2012; Kriks et al., 2011). La greffe de ces neurones dans des modèles animaux immunodéprimés pour la MP (souris, rats et singes) montrent une bonne survie du greffon et une restauration complète des comportements moteurs anormaux induits dans ces modèles (rotation asymétrique, akinésie des membres antérieurs) et surtout l'absence de prolifération conduisant à la formation de tumeurs. En 2014, l'association mondiale (GForce-PD (http://www.gforce-pd.com) de chercheurs et cliniciens provenant d'Europe, des Etats-Unis et du Japon se sont réunis pour coordonner les études précliniques sur la base de thérapies cellulaires pour le traitement de la MP (Barker et al., 2015). Les premières

études cliniques débuteront en 2018, et ouvriront peut-être une perspective nouvelle de thérapie de remplacement pour la MP.

4 Etiologie

Les causes réelles de la MP restent aujourd'hui très peu connues. L'âge est le facteur de risque le plus important pour le développement de la maladie de Parkinson. La prévalence et l'incidence augmentent presque exponentiellement avec l'âge et culminent après 70 ans. Avec le vieillissement de la population et l'augmentation de l'espérance de vie dans le monde, le nombre de personnes atteintes de la maladie de Parkinson devrait augmenter de plus de 50% d'ici 2030. (Voir paragraphe 1 Epidémiologie)

Hormis les rares formes familiales de la MP causées par des mutations de gènes connus, qui touchent moins de 10% des patients, il semble que les formes sporadiques majoritaires soient multifactorielles, associant des facteurs de risques environnementaux et une prédisposition génétique.

4-1 Facteurs environnementaux

Dans les années 1980, l'identification de la neurotoxine, MPTP (1-méthyl-4-phényl-1,2,3,6-tétrahydropyridine) suite à l'étude de jeunes toxicomanes ayant développé un syndrome parkinsonien aigu (Ryskalin et al., 2018), a révélé l'influence de l'environnement sur l'apparition de la MP (Langston and Ballard, 1983).

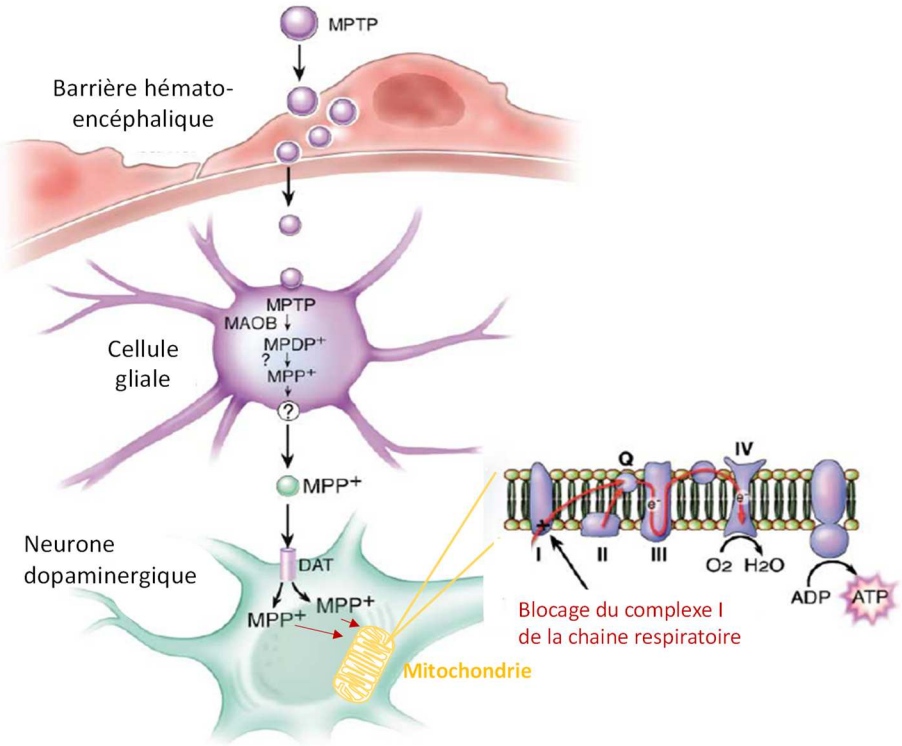


Figure 7 : Représentation schématique du métabolisme du MPTP (adapté de Dauer, 2003).
MPTP : 1-méthyl-4-phényl-1,2,3,6-tétrahydropyridine, MAOB : monoamine oxidase B, MPP+: 1-methyl-4-phenylpyridindisium, DAT : transporteur de la dopamine, complexes mitochondriaux I, II, III et IV, Q : coenzyme Q, O_2 : oxygène, H_2O : eau, ADP : Adénosine diphosphate, ATP : Adénosine triphosphate.

La MPTP est métabolisée en MPP^+ (1-méthyl-4-phenylpyridinium) par la monoamine oxydase B (MAOB), pour être captée par le transporteur de la dopamine (DAT) dans les neurones DA de la SN*pc* et interférer avec le complexe I de la chaine respiratoire de la mitochondrie entrainant l'augmentation d'un stress

oxydatif et une diminution de la synthèse d'ATP (Adénosine triphosphate) (Figure 7). Les personnes exposées accidentellement à la MPTP subissent une perte aiguë de neurones DA et développent des signes et symptômes moteurs parkinsoniens.
Un autre facteur de risque important dans le développement de la MP est l'exposition aux pesticides agricoles. Par exemple, la roténone et le paraquat présentant des similitudes de structure avec la MPTP, augmenteraient de 3 fois le risque de développer la MP impliquant deux mécanismes : l'altération de la fonction mitochondriale en inhibant le complexe I ou l'augmentation du stress oxydatif (Tanner et al., 2011). La roténone pénètre dans les neurones DA indépendamment de DAT et inhibe directement le complexe I de la chaine respiratoire de la mitochondrie. Chez la souris, il en résulte une dégénérescence spécifique de la voie nigro-striée associée à la production de corps de Lewy provoquant une bradykinésie et une rigidité (Betarbet et al., 2000). Le paraquat est non spécifique des neurones DA mais entraine néanmoins une perte neuronale pouvant s'accompagner d'une accumulation d'α-synucléine (Manning-Bog et al., 2002). Récemment, plusieurs études épidémiologiques françaises montrent que l'incidence de la MP est plus élevée dans les zones rurales et plus particulièrement dans les régions viticoles (Kab et al., 2017; van der Mark et al., 2012; Van Maele-Fabry et al., 2012). L'influence de l'exposition aux pesticides domestiques est à confirmer mais l'incidence de la MP chez les non-agriculteurs dans les régions hautement agricoles a tendance à augmenter. Ces populations pourraient être plus spécifiquement ciblées pour des études futures sur l'influence de l'exposition aux pesticides dans l'étiologie de la MP.

L'exposition prolongée aux solvants fait aussi partie des risques de développer la MP. Leur propriété lipophile leur permet de pénétrer dans les systèmes nerveux central et périphérique. Une méta-analyse de 16 études épidémiologiques a montré l'association de la MP à l'utilisation de solvants et l'exposition au trichloréthylène d'un modèle de rongeur a conduit au développement des caractéristiques pathologiques de la MP (Gash et al., 2008; Pezzoli and Cereda, 2013).

Parmi les principaux facteurs environnementaux, les traumatismes crâniens sont associés à un risque accru de la MP (Goldman et al., 2006). La gravité de ces blessures à la tête ou une inconscience de plus de 5 minutes pouvant être provoquée par celles-ci, augmenteraient de 2 fois le risque de développer la MP (Bower et al., 2003; Harris et al., 2013). Chez les personnes présentant une répétition nucléotidique (*Rep1*) dans le promoteur de *SNCA* et dont les taux d'α-synucléine sont élevés, un traumatisme crânien pourrait initier ou accélérer le processus de neurodégénérescence (Goldman et al., 2012). Plusieurs hypothèses concernant les mécanismes impliqués sont avancées : une réaction inflammatoire résultant du processus de réparation pourrait être délétère pour le cerveau, ou la barrière hémato-encéphalique perturbée deviendrait perméable aux neurotoxines , (Bower et al., 2003; Fang et al., 2012; Goldman et al., 2012).

Par ailleurs, le tabagisme et la consommation de café réduiraient le risque de développer la MP (Elbaz et al., 2016), ce qui renforce le concept selon lequel certains facteurs environnementaux modifient la susceptibilité à la maladie de Parkinson.

4-2 Facteurs génétiques

Les avancées de la génétique de ces 20 dernières années, ont permis l'identification de mutations de gènes causant les formes familiales de la MP. Ces formes monogéniques sont rares cependant car seulement 5 à 10% des patients atteints de la MP sont porteurs de ces mutations. À ce jour, au moins 23 gènes pathogènes sont identifiés et désignés comme gènes causant la maladie de Parkinson par le

Comité de nomenclature génétique de l'Organisation du Génome Humain (HGNC) (Tableau 1) (Deng et al., 2018).

Ces gènes codent tous des protéines, leurs mutations entrainent des malformations protéiques et de ce fait un dysfonctionnement du mécanisme dans lequel elles interviennent, provoquant alors des symptômes précoces ou tardifs de la MP associés ou non à la formation de corps de Lewy. Ces gènes mutés se transmettent au sein d'une famille soit par transmission autosomique dominante (un seul allèle muté transmis par un parent est suffisant à la manifestation de la maladie), soit par transmission autosomique récessive (les deux allèles mutés issus des deux parents sont nécessaires pour conduire à la maladie) :

Transmission	Locus	Localisation	Gène	Nom complet	Début de la MP	Corps de Lewy
Autosomique Dominante	***PARK1 PARK4***	4q22.1	*SNCA*	synuclein-alpha	Précoce ou tardive	C
	PARK3	2p13	*PARK3*	Parkinson disease 3	Tardive	NC
	PARK5	4p13	*UCHL1*	ubiquitin C-terminal hydrolase L1	Précoce ou tardive	NC
	PARK8	12q12	*LRRK2*	leucine rich repeat kinase 2	Tardive	C
	PARK11	2q37.1	*GIGYF2*	GRB10 interacting GYF protein 2	Tardive	NC
	PARK13	2p13.1	*HTRA2*	HtrA serine peptidase 2	Précoce ou tardive	NC
	PARK17	16q11.2	*VPS35*	VPS35, retromer complex component	Tardive	NC
	PARK18	3q27.1	*EIF4G1*	eukaryotic translation initiation factor 4 gamma 1	Tardive	NC
	PARK21	20p13	*TMEM230*	transmembrane protein 230	Précoce ou tardive	C
	PARK22	7p11.2	*CHCHD2*	coiled-coil-helix-coiled-coil-helix domain containing 2	Précoce ou tardive	NC
		11p15.4	*RIC3*	RIC3 acetylcholine receptor chaperone	Précoce ou tardive	NC
Autosomique Récessive	***PARK2***	6q26	*PRKN*	parkin RBR E3 ubiquitin protein ligase	Précoce	NC
	PARK6	1p36	*PINK1*	PTEN induced putative kinase 1	Précoce	NC
	PARK7	1p36.23	*PARK7*	parkinsonism associated deglycase	Précoce	NC
	PARK9	1p36.13	*ATP13A2*	ATPase 13A2	Précoce	NC
	PARK14	22q13.1	*PLA2G6*	phospholipase A2 group VI	Précoce	NC
	PARK15	22q12.3	*FBXO7*	F-box protein 7	Précoce	NC
	PARK19	1p31.3	*DNAJC6*	DnaJ heat shock protein family (Hsp40) member C6	Précoce	NC
	PARK20	21q22.1	*SYNJ1*	synaptojanin 1	Précoce	NC
	PARK23	15q22.2	*VPS13C*	vacuolar protein sorting 13 homolog C	Précoce	NC
liée au X	***PARK12***	Xq21-q25	*PARK12*	Parkinson disease 12	Tardive	NC
Non déterminée	***PARK10***	1p32	*PARK10*	Parkinson disease 10	Tardive	NC
	PARK16	1q32	*PARK16*	Parkinson disease 16	Tardive	NC

C= confirmé, NC= non confirmé

Tableau 1 : Gènes impliqués dans les formes familiales de la maladie de Parkinson

- *SNCA*, un gène à transmission autosomique dominante (*locus PARK1* et *PARK4*) :

Le gène *SNCA* code la protéine α-synucléine qui est un composant majeur des corps de Lewy. Ce fut le premier gène associé au phénotype parkinsonien suite à la découverte de la mutation A53T dans une

grande famille italienne (Polymeropoulos et al., 1997). Cinq autres mutations ponctuelles (*PARK1*) ont ensuite été décrites mais restent rares : A30P, E46K, H50Q, G51D, A53E (Deng et al., 2018). Toutes conduisent à des variations faux-sens de la protéine α-synucléine. Il existe également des multiplications du *locus* entier (*PARK4*) induisant une surexpression de l'α-synucléine (Figure 8) (Ryskalin et al., 2018). Les patients avec des multiplications, en particulier ceux avec des duplications homozygotes et jusqu'à quatre copies, présentent des symptômes plus précoces et sévères que ceux avec une seule mutation de *SNCA* (Konno et al., 2016; Lesage and Brice, 2009). En condition physiologique, l'α-synucléine est une petite protéine soluble (140 aa) nativement dépliée. Elle est retrouvée dans le système nerveux central et périphérique, englobant un grand nombre de régions du cerveau (néocortex, substance noire, hippocampe, bulbe olfactif, striatum, thalamus et cervelet) (Ryskalin et al., 2018). Cette protéine est principalement exprimée dans les sites présynaptiques où elle se lie aux membranes (Jakes et al., 1994; Maroteaux et al., 1988). Le rôle de l'α-synucléine n'est pas encore bien compris mais son expression prédominante au niveau des terminaisons pré-synaptiques suggère qu'elle interviendrait dans la régulation de la dynamique des vésicules synaptiques, la fonction mitochondriale et le trafic intracellulaire, et pourrait être une protéine chaperonne (Burre, 2015; Nemani et al., 2010).

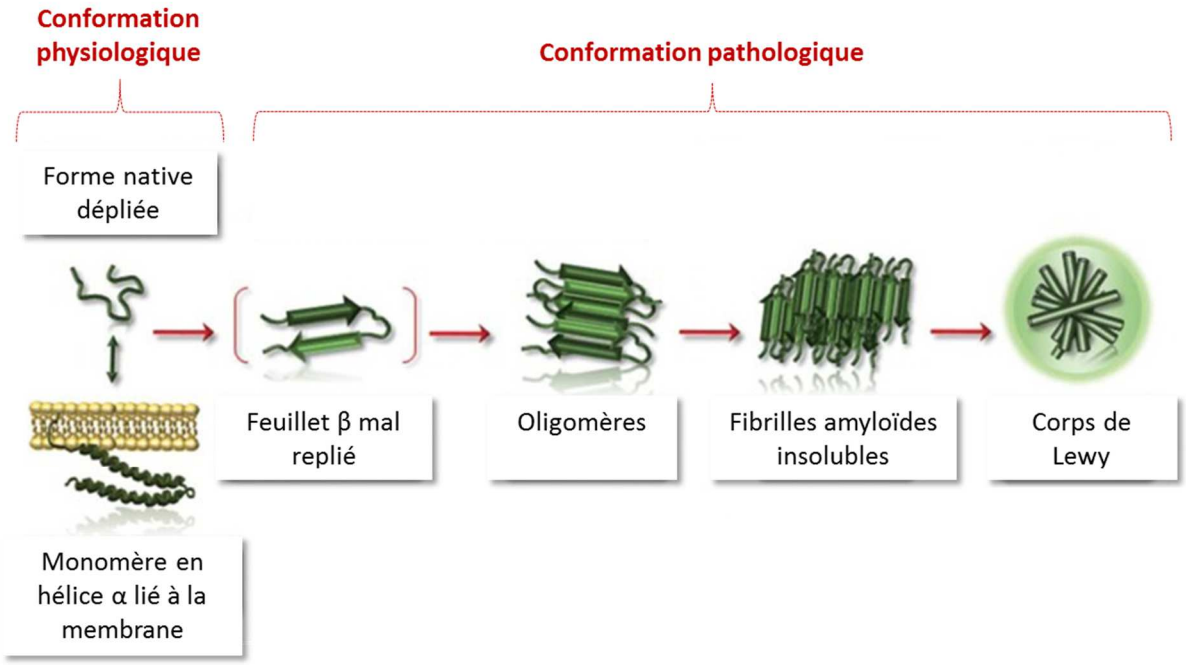


Figure 8 : Représentation schématique du mécanisme d'agrégation de l'α-synucléine (adapté de Ryskalin, 2018).

- ***PRKN*, un gène à transmission autosomique récessive (*locus PARK2*) :**

Découvertes en 1997, les délétions des exons 3 et 7 du gène *PRKN* codant la protéine Parkine furent associées aux formes juvéniles autosomiques récessives de la MP (Kitada et al., 1998; Matsumine et al., 1997). Les mutations du gènes *PRKN* sont à l'origine de 80% des formes autosomiques récessives chez les patients de moins de 20 ans et diminuent significativement lorsque l'apparition de la MP survient après 50 ans (Corti et al., 2011). Plus de 268 mutations de ce gène sont recensées dans la base de données des mutations génétiques humaines de l'Institut Médical de Cardiff, incluant des délétions ou des multiplications d'exons, de petites mutations insertion/délétion et des mutations faux-sens (Deng et al., 2018; Nuytemans et al., 2010). La Parkine est une E3 ubiquitine ligase qui s'associe avec les

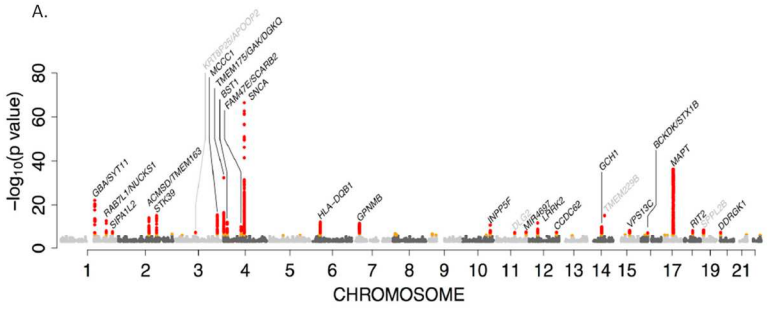
enzymes d'activation et de conjugaison de l'ubiquitine (E1, E2) pour déclencher la dégradation de protéines cibles par le protéasome (van der Merwe et al., 2015). Les mutations du gène *PRKN* provoqueraient un dysfonctionnement de l'ubiquitine E3 ligase, empêchant la dégradation des protéines cibles par le protéasome, entraînant une accumulation de celles-ci et causant une neurotoxicité principalement dans la substance noire (Kalinderi et al., 2016).

- *PINK1*, un autre gène à transmission autosomique récessive (*locus PARK6*) :

Grâce à l'étude du génome d'une grande famille sicilienne, les mutations du gène *PINK1* codant la protéine PTEN Induced Putative Kinase 1 (PINK1) furent également associées aux formes juvéniles autosomiques récessives de la MP (Valente et al., 2001). Il s'agit d'un des gènes les plus fréquemment mutés après le gène *PRKN* pour ces formes de la MP. Plus de 111 mutations ponctuelles ont été recensées dans la MP, entrainant des changements du cadre de lecture, ou tronquant la protéine PINK1. La protéine PINK1 est une sérine/thréonine kinase localisée dans la mitochondrie. Son rôle serait de protéger les cellules contre le dysfonctionnement mitochondrial induit par un stress en initiant la mitophagie (Lazarou et al., 2015). PINK1 phosphoryle l'ubiquitine pour activer l'ubiquitine ligase Parkine qui assemble des chaînes d'ubiquitine sur les protéines de la membrane externe de la mitochondrie « défectueuse », servant ainsi de signal pour la mitophagie. Ainsi les protéines PINK1 et Parkine jouent un rôle clé dans le maintien de l'homéostasie mitochondriale et assurent le contrôle qualité de la mitochondrie. Les mutations des gènes *PRKN* et *PINK1* démontrent encore une fois l'implication d'un dysfonctionnement mitochondrial dans la MP.

À l'exception des formes héréditaires rares de la MP décrites précédemment, 90% des patients souffrent de formes sporadiques difficiles à expliquer. Les techniques récentes d'analyse du génome sur de larges cohortes de patients (GWAS = Genome Wide Association Study) ont permis de démontrer le rôle déterminant des facteurs génétiques dans l'étiologie de cette pathologie, même si nombre d'entre eux n'ont pas encore été identifiés (Mullin and Schapira, 2015; Verstraeten et al., 2015).

A.



B.

SNP	Chr	Position (pb)	Gène(s) voisin(s)
rs35749011	1	155,135,036	GBA/SYT11
rs823118	1	205,723,572	RAB7L1/NUCKS1
rs10797576	1	232,664,611	SIPA1L2
rs6430538	2	135,539,967	ACMSD/TMEM163
rs1474055	2	169,110,394	STK39
rs115185635	3	87,520,857	KRT8P25/APOOP2
rs12637471	3	182,762,437	MCCC1
rs34311866	4	951,947	TMEM175/GAK/DGKQ
rs11724635	4	15,737,101	BST1
rs6812193	4	77,198,986	FAM47E/SCARB2
rs356182	4	90,626,111	SNCA
rs9275326	6	32,666,660	HLA-DQB1
rs199347	7	23,293,746	GPNMB
rs117896735	10	121,536,327	INPP5F
rs3793947	11	83,544,472	DLG2
rs329648	11	133,765,367	MIR4697
rs76904798	12	40,614,434	LRRK2
rs11060180	12	123,303,586	CCDC62
rs11158026	14	55,348,869	GCH1
rs1555399	14	67,984,370	TMEM229B
rs2414739	15	61,994,134	VPS13C
rs14235	16	31,121,793	BCKDK/ STX1B
rs17649553	17	43,994,648	MAPT
rs12456492	18	40,673,380	RIT2
rs62120679	19	2,363,319	SPPL2B
rs8118008	20	3,168,166	DDRGK1

Figure 9 : Découverte de 26 loci à risques associés à la maladie de Parkinson (adapté de Nalls, 2014).

(A) Le graphique représente les résultats d'une méta-analyse des données SNP (=polymorphisme nucléotidique) à l'échelle du génome sur la base d'un seuil de valeur de probabilité (p-value) supérieur à 5.10^{-8} (points oranges et rouges) sur les différents chromosomes. (B) Le tableau recense les 26 SNP principaux, leur position sur le chromosome (Chr) et le(s) gène(s) voisin(s) le(s) plus proche(s).

Grâce aux avancées technologiques dans le domaine de la génomique, la communauté scientifique a été littéralement inondée de données concernant des polymorphismes nucléotidiques (ou SNP = single nucleotide polymorphism) associés à la quasi-totalité des grandes pathologies humaines. Pour les formes sporadiques de la MP, une grande étude d'association à l'échelle du génome de 13 708 patients parkinsoniens et 95 282 contrôles a révélé 26 *loci* à risques dont 6 nouveaux *loci* : *SIPA1L2, INPP5F, MIR4697, GCH1, VPS13C* et *DDRGK1* (Figure 9) (Nalls et al., 2014).

En résumé, l'étude des formes héréditaires de la MP ont fourni des indices cruciaux sur les mécanismes sous-jacents à la neuropathologie de la maladie de Parkinson. Certaines des protéines codées par les gènes associés à la MP sont impliquées dans un ensemble de voies moléculaires qui, lorsqu'elles sont altérées, peuvent déclencher une neuropathologie. En outre, de grandes études d'association à l'échelle du génome confirment que certains de ces gènes sont également affectés dans les formes sporadiques de la MP. Aujourd'hui, les principaux processus biologiques mis en cause dans la mort neuronale et cités précédemment sont : la dérégulation des mécanismes de dégradation des protéines, le dysfonctionnement mitochondrial et le stress oxydatif, le dysfonctionnement du transport axonal et la neuro-inflammation.

Cependant et de façon intéressante, une proportion importante des régions identifiées comme facteurs de risques par ces grandes études d'association sont localisées dans des régions génomiques non-codantes (Brodie et al., 2016). Dans ces régions, une des classes d'éléments non codants se distingue ces dernières années, les longs ARN non-codants (lncARN).

B Les longs ARN non codants

Le transcriptome représente l'ensemble des ARN codant des protéines et des ARN non codants (ncARN) issus de la transcription du génome d'une cellule ou d'un tissu. Depuis la fin des années 50, il est admis que des mécanismes essentiels de la machinerie cellulaire sont assurés par des ncARN comme les ARN de transfert (ARNt) ou les ARN ribosomaux (ARNr) qui sont des ncARN très abondants ayant principalement des rôles structurels et régulateurs dans la machinerie traductionnelle. Grâce au projet sur le génome humain (ENCODE) et l'arrivée du séquençage de nouvelle génération à haut débit, une grande partie de l'ADN (90%) s'est révélée transcrite (Carninci et al., 2005; Consortium, 2012) générant un large répertoire de molécules d'ARN dont la majorité sont des long ARN non codants (lncARN) (Guttman et al., 2009). Sur la totalité du transcriptome humain, seulement 28% des transcrits sont annotés dans les bases de données, 36% d'entre eux correspondent à des gènes codants et 39% à des gènes non codants (Figure 10). Ces derniers se répartissent selon leur taille en deux catégories : les petits ARN non codants de taille inférieure à 200 nucléotides et longs ARN non codants (66%) de taille supérieure à 200 nucléotides (Figure 10). Les premiers petits ARN non codants identifiés sont les petits ARN nucléolaires (snoRNA) impliqués dans la maturation des ARNr et les petits ARN nucléaires (snRNA) impliqués dans l'épissage des ARN messagers (ARNm). Puis d'autres intervenant comme régulateurs de l'expression des gènes par ARN interférence sont découverts comme les miRNA (micro ARN), les siRNA (small interfering RNA) et les piRNA (Piwi-interacting RNA) (Castel and Martienssen, 2013).

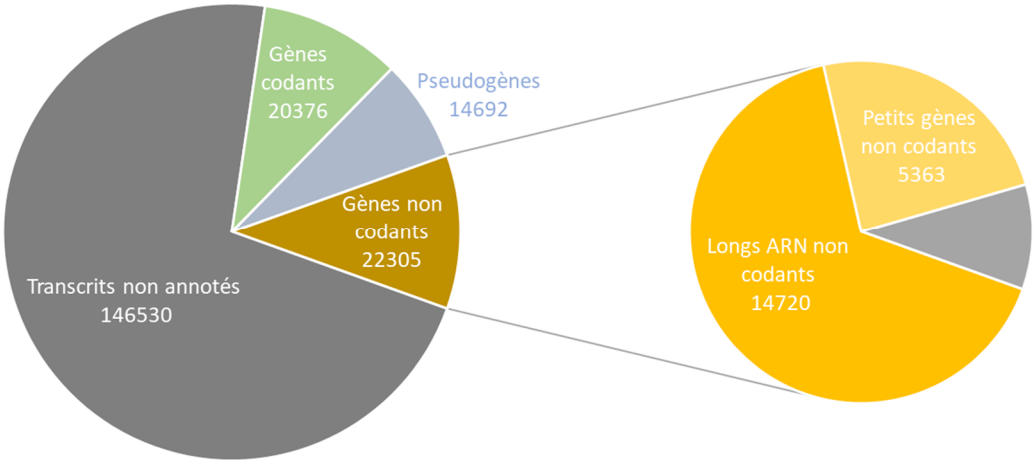


Figure 10 : Répartition des transcrits du génome humain.
Les gènes annotés se divisent en trois catégories : les gènes codant des protéines, les pseudogènes (= anciens gènes codants inactifs) et gènes non codants. (Diagrammes réalisés à partir des données de la dernière version de GENCODE = version 28 de janvier 2018, Genome de Référence Humain GRCh38).

Enfin, une dernière catégorie de longs transcrits fut découverte dans les années 1990 avec les premiers lncARN, tels que *H19* impliqué dans l'empreinte parentale et *XIST* (*X inactive specific transcript*) responsable de l'inactivation du chromosome X maternel (Bartolomei et al., 1991; Brown et al., 1991). Ces deux exemples démontrent la complexité du transcriptome non codant et sa pertinence dans les processus biologiques. Depuis les années 2000, de nombreux lncARN sont identifiés et leur implication dans les processus cellulaires fondamentaux, comme la régulation de l'expression des gènes, la transcription et la traduction ont mis en évidence leur rôle central dans l'homéostasie cellulaire.

Cependant, le manque de systèmes d'annotation systématiques et uniformes engendre des noms différents pour un même lncARN rendant la concordance des résultats scientifiques difficile. Aujourd'hui, nombre de lncARN n'ont pas encore été annotés et très peu ont été caractérisés fonctionnellement.

1 Définition d'un long ARN non-codant

1-1 Caractéristiques des longs transcrits non codants

Nos connaissances sur les longs ARNs non codants (lncARN) restent partielles contrairement à celles sur les ncARN de petites tailles, bien qu'ils soient de plus en plus étudiés. Selon le Comité de nomenclature génétique de l'Organisation du Génome Humain (HGNC), les principaux critères qui définissent un long ARN non codant sont :

- la taille, supérieure à 200 nucléotides et jusqu'à 100 kb (kilobases),
- l'incapacité à coder une protéine,
- les lncARN présentent dans leur grande majorité des similitudes avec les ARNm (Guttman et al., 2009; Jarroux et al., 2017):
  - la plupart des lncARN eucaryotes sont transcrits par l'ARN polymérase II et requièrent le recrutement de facteurs de transcription,
  - ils subissent des modifications post-transcriptionnelles augmentant leur stabilité,
  - certains présentent une coiffe 7mGppp à leur extrémité 5' les rendant moins vulnérables aux exonucléases et/ou favorisant leur exportation du noyau au cytoplasme,
  - ils sont pour la plupart polyadénylés à leur extrémité 3',
- ils peuvent être mono ou multi-exoniques : une grande partie des lncARN sont composés de 2 exons (42% contre 6% des ARNm) et leur taille est généralement plus longue en moyenne (Derrien et al., 2012)
- ils sont par conséquent soumis à des processus d'épissage : leurs signaux d'épissages sont similaires à ceux des ARNm (Derrien et al., 2011)
- ils possèdent les marques épigénétiques des modifications des histones H3K4me3 et H3K36me3 qui sont des signatures respectivement des sites de démarrage de la transcription (TSS) et d'élongation (Cabili et al., 2011; Guttman et al., 2009) comme pour les gènes codants.

Enfin, les lncARN ont un niveau d'expression généralement plus faible que les ARNm rendant leur étude difficile et ils sont très spécifiques d'un type cellulaire (Cabili et al., 2011; Derrien et al., 2012; Moran et al., 2012). Notre laboratoire en collaboration de l'équipe de Jorge Ferrer, a été l'un des premiers à montrer la spécificité tissulaire des lncARN. Les lncARN de la cellule β pancréatique humaine sont plus spécifiques de ce type cellulaire que ne le sont les ARN codants (Moran et al., 2012). J'ai activement participé à cette étude. Dernièrement notre laboratoire a mis en évidence le répertoire de lncARN des neurones DA issus du mésencéphale ventral d'embryons de souris au stade E14.5. (Gendron et Colace-Sauty et al, 2018 - en révision). En comparant ce répertoire avec un répertoire spécifique de neurones sérotoninergiques du rhombencéphale, ainsi que des données de la littérature, cette étude a permis de mettre en évidence une proportion très importante de longs ARNs non codants exprimés uniquement dans les neurones DA, renforçant ainsi l'idée de la spécificité cellulaire des lncARN. De nombreuses études ont également confirmé cette propriété (Mele et al., 2017; Ulitsky and Bartel, 2013; van Rossum

et al., 2016). Cette particularité rend la découverte des lncARN encore plus difficile et dépendante de l'accessibilité aux cellules d'intérêt.

1-2 ARN codant ou non codant ?

Du fait de leur similitude avec les ARNm, le caractère non codant d'un lncARN nouvellement identifié est un critère essentiel. Le potentiel codant d'un gène repose principalement sur la détermination du cadre de lecture de l'ARN (ou ORF= Open Reading Frame), l'identification des codons d'initiation et de terminaison, et la reconnaissance des jonctions d'épissage. Différentes méthodes existent et permettent de discriminer un ARN codant d'un ARN non codant :

- la méthode CPAT (Coding Potential Assessment Tool) (Wang et al., 2013) : il s'agit d'une méthode d'analyse bio-informatique visant à évaluer le potentiel codant d'un transcrit en se fondant non seulement sur l'identification de la présence de cadres de lecture ouverts mais aussi sur la fréquence d'utilisation des codons (spécifiques d'une espèce) et l'enchainement des codons qui n'est pas équiprobable dans les transcrits codants. En d'autre termes, dans la séquence protéique d'une espèce donnée, après un acide aminé particulier la probabilité de trouver n'importe quel autre acide aminé n'est pas équivalente, certains sont très fréquents d'autres très rares. De fait, il est possible de prédire un score codant à partir de la suite des codons retrouvés dans une phase ouverte de lecture. Cette méthode permet de calculer un score global de prédiction du caractère codant d'un ARN et d'exclure le potentiel codant d'un transcrit.
- le « ribosome profiling » ou Ribo-Seq (Guttman et al., 2013; Ingolia, 2014) : il s'agit d'une méthode basée sur la détection des ARN associés aux ribosomes et donc susceptibles d'être traduits en protéines. Cependant ce seul paramètre n'est pas suffisant pour affirmer qu'un transcrit est codant. Un autre paramètre basé sur une propriété de la traduction, le relargage du ribosome au codon stop, permet cette discrimination. En effet, les taux de ribosomes associés aux lncARN sont similaires à ceux trouvés dans les régions 5'UTR (Untranslated Transcribed Region) des ARN codants mais il n'y a pas de signal de relargage des ribosomes au codon stop pour les lncARN contrairement aux ARNm.
- la spectrométrie de masse des protéines (MS) : il s'agit d'une méthode basée sur la détection des produits peptidiques issus de la traduction par mesure de leur masse. Les phases ouvertes de lectures des lncARN permettent de prédire la séquence de polypeptides. En MS, ces polypeptides ne sont pas retrouvés ou seulement en très faible abondance. En effet, dans le cadre du projet ENCODE, les niveaux d'expression des peptides provenant potentiellement de lncARN dans deux lignées cellulaires était 13 à 20 fois moins élevés que ceux provenant des ARNm et 92% des lncARN annotés ne sont pas traduits (Banfai et al., 2012).

Pour conclure, les lncARN montrent certaines similarités structurelles avec les ARNm, cependant ils ont un faible potentiel de traduction malgré leur association avec des ribosomes. Bien qu'il soit possible que les faibles niveaux de détections des peptides issus de la traduction de lncARN soient dus aux limitations de la MS pour les peptides inférieurs à 50 aa, il semble que ces peptides soient très instables et ne s'accumulent pas à des taux conséquents dans les cellules de mammifères (Housman and Ulitsky, 2016).

2 Classification des lncARN

2-1 Evolution des méthodes d'analyse des ARN non codants

Le premier catalogue à grande échelle de transcrits supposés non codants provient du projet FANTOM (Carninci et al., 2005; Okazaki et al., 2002). Grâce à la technique de séquençage Sanger d'ADNc (ADN complémentaire) clonés, plus de 34 000 ARN non codants exprimés dans différents tissus chez la souris ont pu être identifié (Ravasi et al., 2006). Des études ultérieures ont permis d'affiner les catalogues de lncARN chez la souris et chez l'homme, constituant les annotations actuelles de lncARN des bases de données RefSeq et Ensembl (Derrien et al., 2012; Pruitt et al., 2012). Parallèlement, des puces à ADN dites « Tiling microarray » ont été utilisées pour détecter de nouvelles régions transcrites (Bertone et al., 2004; Guttman et al., 2009; Khalil et al., 2009), potentiellement plus sensibles que le clonage mais toutefois limitées car il est difficile de définir des jonctions d'épissage et de relier les régions transcrites entre elles (Agarwal et al., 2010). Plus récemment, le séquençage à haut débit de millions de fragments courts d'ARN d'environ 100 nucléotides (RNA-seq) permet de reconstruire les transcrits même très faiblement exprimés à l'aide d'un génome de référence (Cabili et al., 2011; Pauli et al., 2012; Trapnell et al., 2010). De plus, même des transcrits plus rares peuvent être spécifiquement enrichis en utilisant des méthodes de capture en amont du séquençage (Mercer et al., 2011). L'ARN CaptureSeq consiste à hybrider des ADNc spécifiques de régions génomiques d'intérêt, puis les éluer et les séquencer, permettant ainsi d'enrichir un échantillon avec des ARN « captés » avant leur séquençage.

2-2 Différentes classifications

La classification des lncARN est aujourd'hui encore en constante évolution, agrémentée par la découverte de nouveaux lncARN en fonction de l'affinement des méthodes d'études du génome à grande échelle et par la clarification de leur fonction. Plusieurs classifications des lncARN sont proposées, sachant que cette liste n'est pas exhaustive (Jarroux et al., 2017; St Laurent et al., 2015) :

- Selon leur taille : les long ARN non codants (lncARN) de taille supérieure à 200pb se distinguent des très long ARN non codants (vlincARN et macro lncARN) de taille supérieure à 10kb (Guenzl and Barlow, 2012; St Laurent et al., 2016).
- Selon leur localisation dans une région particulière du génome : certains sont par exemple des pseudogènes qui correspondent à d'anciens gènes qui ont perdu leur potentiel codant au cours de l'évolution. Les lncARN dérivés de pseudogènes peuvent réguler l'expression du gène codant en agissant comme des éponges de miRNA, comme c'est le cas de *PTENP1*, un lncARN pseudogène dérivé du gène suppresseur de tumeur *PTEN*. La présence de *PTENP1* régulerait l'expression de *PTEN* par compétition au niveau de la fixation de miRNA sur des séquences identiques aux deux gènes, et ainsi jouerait un rôle protecteur contre les cancers. Les expressions de ces deux gènes ont été retrouvées diminuées dans les cancers de la prostate et du colon (Poliseno et al., 2010). D'autres par association aux télomères des chromosomes, les TERRA (telomeric repeat containing RNA) favorisent la réponse aux dommages de l'ADN télomérique et par conséquent, retardent la sénescence (Porro et al., 2010).
- Selon leur association avec des éléments génomiques connus : certains lncARN sont associés à des enhancers, il s'agit de longs transcrits de séquences enhancers appelés elncARN. Ils régulent l'expression de gènes avoisinants en agissant en *cis* sur leurs enhancers pour recruter des facteurs de transcription. Par exemple, le elncARN *Evf2* possède les doubles fonctions d'activateur ou de répresseur des gènes à proximité selon les protéines avec lesquelles il s'associe et il joue un rôle critique dans le développement du cerveau (Bond et al., 2009).

D'autres lncARN sont associés à des promoteurs et sont appelés des PALR (Promoter-associated long RNA). Par exemple, des dommages de l'ADN induisent l'expression du PALR du gène *CCND1* (Cycline D1) permettant la répression de l'expression des gènes adjacents (Song et al., 2012).

- Selon leur localisation subcellulaire : la majorité des lncARN sont retrouvés dans le noyau mais certains sont localisés dans le cytoplasme ou encore dans les mitochondries (voir paragraphe Localisation cellulaire)
- Selon leur fonction : de nombreuses fonctions de régulations sont attribuées au lncARN et sont détaillées dans le paragraphe Mécanismes d'action.
- Selon leur organisation chromosomique par rapport aux gènes adjacents : il s'agit d'une classification en fonction de la localisation chromosomique des lncARN qui prend en compte leur position par rapport au gène codant annoté le plus proche. A partir de cette classification, 6 types de lncARN sont définis dans la figure suivante (Figure 11) (Moran et al., 2012).

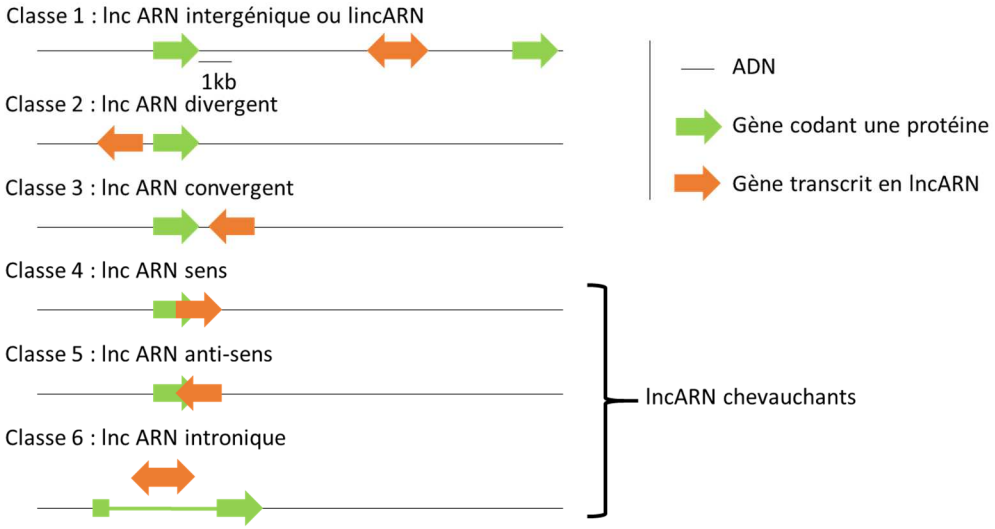


Figure 11 : Organisation chromosomique des lncARN par rapport au gène le plus proche (adapté de Moran et al., 2012).
La séquence d'ADN est représentée par une ligne noire continue, les gènes codants des protéines et les gènes transcrits en lncARN sont respectivement représentés en vert et orange. Le sens de transcription est représenté par une flèche, lorsqu'il y a une double flèche, le lncARN peut être orienté dans les deux sens. 1kb (kilobase) est la distance minimale entre le lncARN et le gène voisin, choisie arbitrairement pour définir la classe de lncARN intergénique.

Classe 1 = lncARN intergénique (LincARN) situé à plus de 1 kb du gène codant le plus proche indépendamment du sens de transcription de celui-ci. La distance minimale de 1kb entre le lncARN et le gène adjacent correspond à la taille moyenne d'un promoteur proximal et a été choisi arbitrairement pour définir la classe de lncARN intergénique (Guttman et al., 2009; Moran et al., 2012). Cette classe a été définie par Guttman et ses collègues en combinant les données du transcriptome obtenues par Tiling Array et la localisation des événements du démarrage de la transcription (TSS) et de l'élongation sur le génome. Ces deux sites sont repérables grâce à la cartographie des marques épigénétiques des modifications des histones H3K4me3 (signature du TSS) et H3K36me3 (signature du site d'élongation). Les LincARN ont été les plus étudiés au niveau de leur expression et de leur conservation parmi les espèces mais aussi de leur implication dans divers mécanismes : comme la régulation transcriptionnelle en *cis* ou en *trans,* la régulation de l'épissage, le contrôle de la traduction, ou encore la régulation par des modifications post-transcriptionnelles (Guttman et al., 2009; Ulitsky and Bartel, 2013).

Les autres classes tiennent compte de l'orientation de la transcription du gène codant le plus proche.

Classe 2 = lncARN divergent. Le sens de sa transcription diverge du gène codant le plus proche (brin d'ADN opposé) et se situe à moins d'1kb de celui-ci sans toutefois le chevaucher.

Classe 3 = lncARN convergent. Le sens de sa transcription converge vers le gène codant le plus proche (brin d'ADN opposé) et se situe à moins d'1kb de celui-ci sans toutefois le chevaucher.

Classe 4 et 5 = lncARN chevauchant en partie ou entièrement la séquence d'un autre gène et sa transcription est orientée dans le même sens (4) que celui-ci (même brin d'ADN) ou en sens inverse (brin d'ADN opposé) (5). Certains lncARN anti-sens (lncARN-AS) peuvent réguler l'expression du gène qu'ils chevauchent comme c'est le cas de MAPT-AS1 qui régule l'expression de MAPT dans la MP (Coupland et al., 2016). Les PALR cités précédemment sont considérés comme des lncARN chevauchants sens car leur séquence chevauche le promoteur et la première région exonique et intronique d'un autre gène. Kapranov et ses collègues ont isolé des lncARN chevauchants sens en séquençant les extrémités 5' d'ADNc codant par RACE-PCR (Rapid amplification of cDNA ends by polymerase chain reaction) (Kapranov et al., 2007).

Classe 6 = lncARN intronique chevauchant un intron d'un autre gène (la double flèche indique que le lncARN peut être orienté dans les deux sens). Cette classe de lncARN a été moins étudiée que les LincARN et les lncARN-AS durant ces dernières années car ils ont longtemps été considérés comme des produits d'épissage alternatif d'ARNm. De plus, ils semblent moins stables que les ARNm et les lncARN-AS (Ayupe et al., 2015).

Cette dernière classification en fonction de la localisation chromosomique des lncARN est particulièrement pertinente pour l'analyse de la conservation des lncARN entre les espèces. En effet, la conservation de l'ordre de groupe de gènes sur un même chromosome chez tous les organismes, informe sur leurs fonctions biologiques cruciales : il s'agit de la notion de synténie qui sera abordée dans le paragraphe ci-après.

3 Conservation

L'analyse de la conservation des lncARN entre les espèces peut fournir des indices importants sur leurs modes d'action et permettre d'identifier des modèles animaux pertinents pour l'étude de lncARN humains. Le séquençage à grande échelle du génome et du transcriptome du monde eucaryote a permis l'analyse comparative de la conservation inter-espèce de la séquence primaire des gènes codants et non codants. Comparé aux gènes codants et aux petits ncARN, les lncRNA sont faiblement conservés au niveau de leur séquence primaire (Ponjavic et al., 2007). Cependant, leurs régions promotrices comprenant le site d'initiation de la transcription, sont en moyenne plus conservées que leurs exons, et dans les mêmes proportions que les gènes codants (Carninci et al., 2005; Derrien et al., 2012; Guttman et al., 2009). L'organisation synténique des lncARN sur les chromosomes constituent un facteur plus adapté à l'étude de l'évolution des lncARN. Une récente étude du transcriptome non codant de 17 espèces différentes a montré, pour plus d'un millier de lncARN, une conservation plus élevée des TSS, de la synténie, et de la spécificité tissulaire. Ces observations suggérent un contrôle par des systèmes de régulation conservés tout en tolérant des changements majeurs dans l'architecture des gènes (Hezroni et al., 2015). De plus, les lncARN les plus conservés fonctionnellement, malgré une conservation de séquence limitée, sont des LincARN jouant des rôles biologiques cruciaux durant le développement embryonnaire.

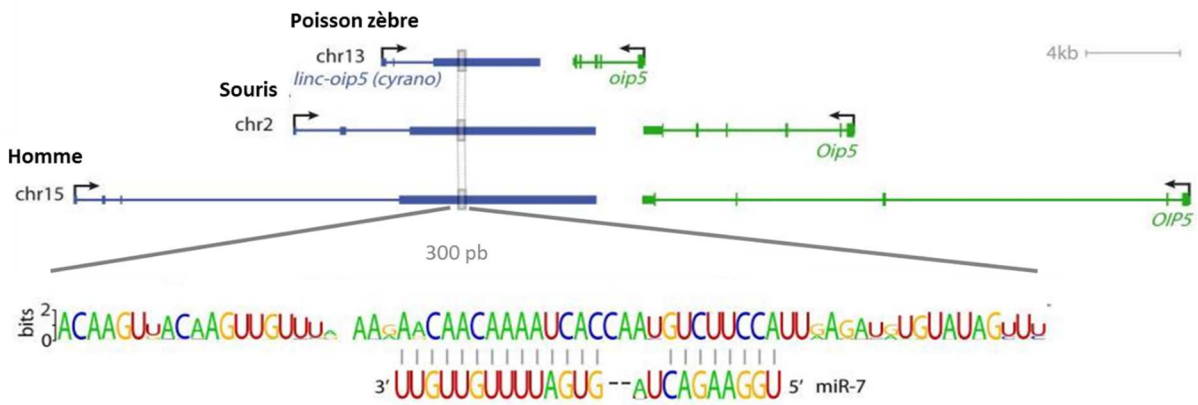


Figure 12 : Exemple de séquence conservée pour le LincARN OIP5-AS1 ou CYRANO (adapté de Ulitsky et al., 2011).
En bleu, le LincARN OIP5-AS1 et en vert le gène codant OIP5 (Opa Interacting Protein 5). Sur les alignements de 45 génomes de vertébrés, 3 sont représentés ici : le poisson zèbre, la souris et l'homme. En gris, la région de 300 paires de bases la plus conservée entre les espèces et en couleurs la séquence d'ARN du poisson zèbre de 67 bases parfaitement conservées dans les 45 génomes incluant la séquence du miR-7 annotée dans miRBase.

Par exemple, le lincARN CYRANO ou OIP5-AS1 partage une structure génétique similaire (2-3 exons courts suivis d'un long exon terminal de 4-8 kb), une même région synténique ainsi qu'une fonction au cours de l'embryogénèse chez le poisson zèbre que conserve les orthologues de mammifères (souris et humain) (Ulitsky et al., 2011). Dans cette étude, la perte de fonction de CYRANO grâce à des oligonucléotides antisens ciblant la région conservée, a provoqué des anomalies du développement qui ont pu être réparées en ajoutant soit le lincARN mature du poisson zèbre, soit son orthologue humain ou de souris. Enfin, sur la séquence primaire du LincARN CYRANO, 300 nucléotides sont très conservés dont 26 nucléotides correspondent à un site de fixation du miR-7 (Figure 12). En effet, une autre étude a montré que CYRANO contrôlait l'action du miR-7 en agissant comme un leurre pour le maintien de l'immortalité des cellules souches embryonnaires (Smith et al., 2017).

Ainsi, l'étude de l'évolution et de la conservation inter-espèces des lncARN peut permettre de prévoir les fonctions biologiques d'autres lncARN et de favoriser des modèles animaux pour leurs études fonctionnelles.

4 Localisation cellulaire

Les ARNnc sont présents chez tous les eucaryotes, des organismes unicellulaires aux mammifères supérieurs. Contrairement aux ARNm matures majoritairement localisés dans le cytoplasme, les lncARN sont retrouvés :

- dans le noyau au niveau de la chromatine ou dans le nucléoplasme : la plupart des lncARN sont nucléaires. Par exemple *NEAT1* (Nuclear Paraspeckle Assembly Transcript 1) joue un rôle dans la formation de « paraspeckles » qui sont des structures nucléaires permettant le recrutement de la machinerie transcriptionnelle (Clemson et al., 2009). L'existence d'un motif AGCCC sur la séquence du LincARN *BORG (*BMP/OP-Responsive Gene), a montré une corrélation directe avec sa présence dans le noyau, suggérant que ce motif peut agir comme un signal de localisation nucléaire (Zhang et al., 2014).
- dans le cytoplasme : certains sont exclusivement cytoplasmiques comme par exemple, *DANCR* (Differentiation Antagonizing Non-Protein Coding RNA) (van Heesch et al., 2014).

- dans les mitochondries : le génome mitochondrial code un certain nombre de lncARN dont *lncND5*, *lncND6* et *lncCyt b*, qui semblent impliqués dans la stabilité des ARN mitochondriaux adjacents (*ND5, ND6 et Cyt b*) (Rackham et al., 2011).

Des changements dans l'environnement, comme un stress oxydatif, peuvent aussi influencer la localisation subcellulaire des lncARN (Giannakakis et al., 2015). Les modifications de compartiments cellulaires existent aussi chez les lncARN, par exemple, les protéines HuR (Hu-Antigen R) et GRSF1 (G-Rich RNA Sequence Binding Factor 1) régissent la localisation cytoplasmique et mitochondriale du lncARN *RMRP* (RNA Component Of Mitochondrial RNA Processing Endoribonuclease) codé par l'ADN nucléaire mais ayant des fonctions clés dans les mitochondries (Noh et al., 2016). Ainsi, il semble évident que la répartition de lncARN au sein d'une cellule influe sur leur fonction.

5 Mécanismes d'action

Comme décrit précédemment, l'étude de la séquence nucléotidique primaire des lncARN ne permet pas de prédire leurs fonctions biologiques. En revanche, leur localisation sur le génome et leur localisation subcellulaire peuvent fournir des indices importants sur leurs modes d'action. Au cours de la dernière décennie, de nombreuses études des lncARN ont mis en lumière leurs divers rôles dans les différents processus biologiques : ils interviennent dans la régulation de l'expression de gènes à travers des mécanismes variés sur l'épigénétique et la transcription jusqu'aux modifications post-transcriptionnelles *via* l'interaction avec de nombreuses molécules (protéines, chromatine, ARN, petits ARN non codants). Cependant, pour la plupart des lncRNA identifiés, leurs fonctions biologiques restent à explorer.

5-1 Mode d'action en cis ou trans

Un lncARN qui a une capacité de régulation en *cis*, peut fonctionner tout en restant attaché à son propre locus contrairement à un ARNm qui doit être exporté et traduit pour fonctionner (Figure 13A). C'est le cas du lncARN *XIST* (*X-Inactive Specific Transcript*), très bien caractérisé et qui joue un rôle important dans l'inactivation du chromosome X maternel chez les mammifères (Brown et al., 1991). *XIST* est un exemple démontrant qu'un lncARN peut jouer un rôle direct dans l'inactivation de grandes régions génomiques en interagissant physiquement avec des complexes protéiques, conduisant à la condensation de la chromatine et à la répression transcriptionnelle d'un chromosome X entier. Chez les femelles, *XIST* est transcrit sur le chromosome X inactif Xi en grande quantité, s'accumule autour de celui-ci et l'inactive spatialement mais aussi, recrute un complexe de ribonucléoprotéines impliquées dans la régulation épigénétique de l'expression des gènes en modifiant les histones (Brown et al., 1991).

D'après Guttman et Rinn, si un lncARN est un *cis*-régulateur, alors les niveaux de l'expression génique d'un gène voisin sur le même allèle seront corrélés avec l'expression du lncARN dans toutes les conditions (Guttman and Rinn, 2012). Peu d'études d'inhibition de lncARN montrent une corrélation d'expression avec un gène voisin sur le même allèle. Ceci indique que, bien que certains lncARN soient des *cis*-régulateurs, la grande majorité de ceux identifiés et caractérisés jusqu'à présent, fonctionne comme des *trans*-régulateurs (Guttman and Rinn, 2012).

Un lncARN qui a une capacité de régulation en *trans*, fonctionne à distance de son site de transcription sur d'autres régions ou chromosomes en recrutant d'autres facteurs nucléoprotéiques (Figure 13B). C'est le cas du lncARN *HOTAIR* (*HOX Transcript Antisens Intergenic RNA*), qui joue un rôle crucial au cours du développement et dans les cancers en guidant des complexes protéiques spécifiques sur sa cible moléculaire. Il est localisé sur le locus des gènes HoxC et réprime, grâce à un mode d'action en *trans*,

les gènes HOXD situés à 40kb de son site de transcription en recrutant et transportant des modulateurs épigénétiques de la chromatine (Rinn et al., 2007).

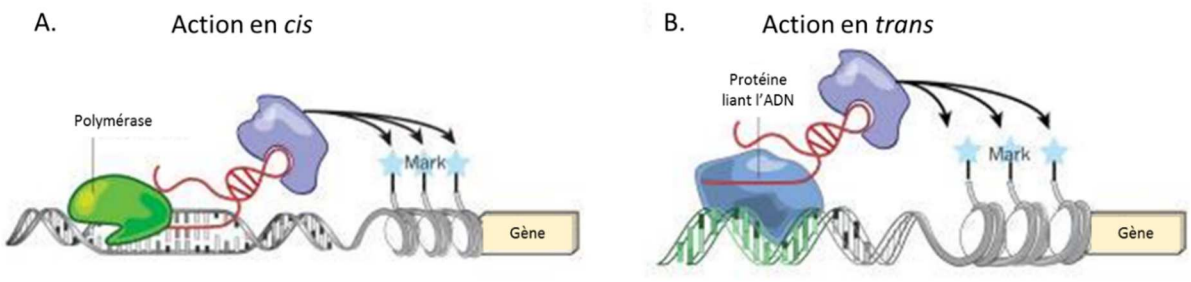
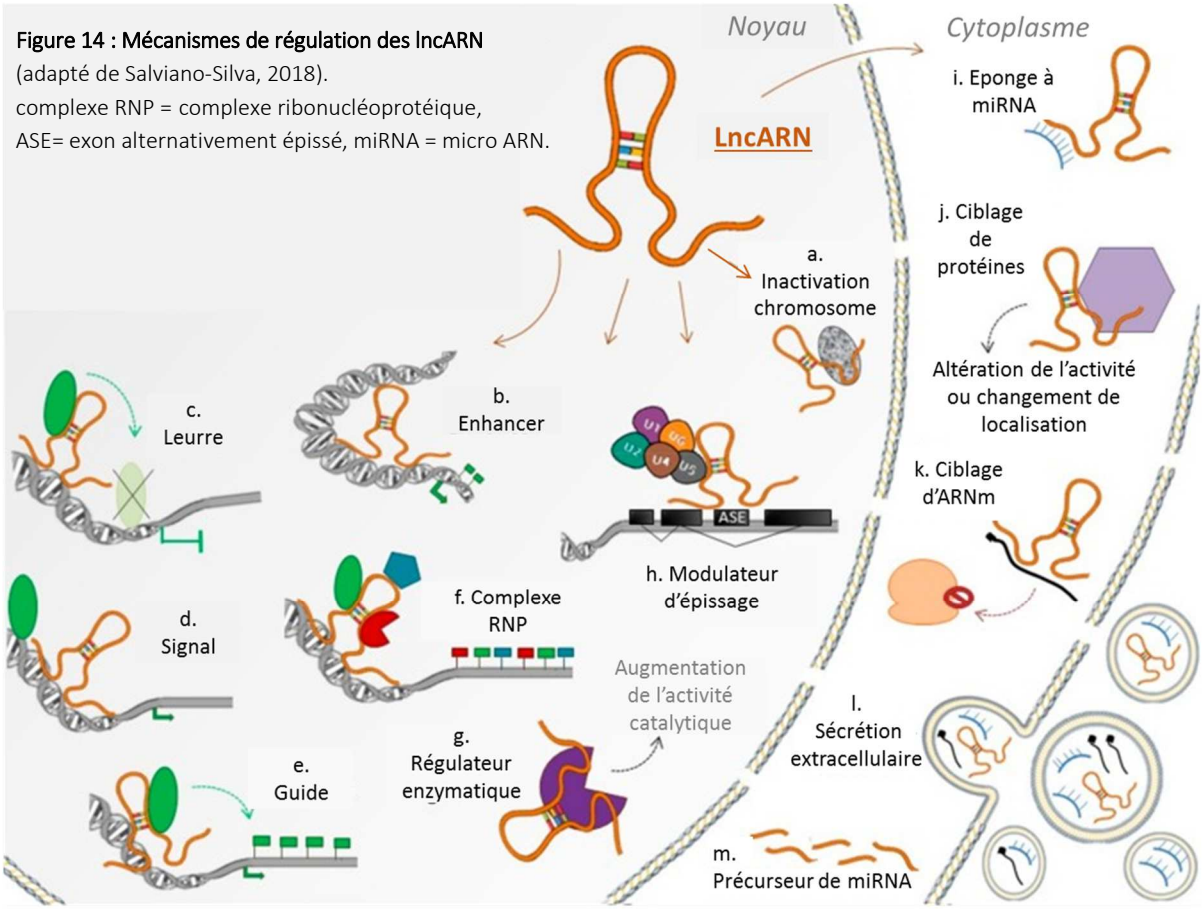


Figure 13 : Mode d'action en cis ou trans des lncARN (adapté de Guttman, 2012)
(A) Un modèle de lncARN qui agit en *cis*-régulation en restant attaché à son site de transcription : l'ARN polymérase (en vert) transcrit un lncARN (rouge), qui peut s'associer aux protéines régulatrices (en violet) pour affecter les régions voisines. (B) Un modèle de lncARN qui agit en trans-régulation : un lncARN transcrit sur un autre *locus* ou un autre chromosome peut s'associer avec des protéines liant l'ADN (bleu) et des protéines régulatrices pour localiser et affecter l'expression de gènes cibles. Mark = modifications épigénétiques.

5-2 Mécanismes moléculaires

Les lncARN sont impliqués dans de nombreux processus biologiques tels que la prolifération cellulaire, le cycle cellulaire, le métabolisme, l'apoptose, la différenciation et le maintien de la pluripotence (Geisler and Coller, 2013; Wang et al., 2011). Il est maintenant établi qu'ils régulent l'expression des gènes, et qu'ils emploient divers mécanismes pour leur fonction (Figure 14) :

Figure 14 : Mécanismes de régulation des lncARN
(adapté de Salviano-Silva, 2018).
complexe RNP = complexe ribonucléoprotéique,
ASE= exon alternativement épissé, miRNA = micro ARN.



a. Les lncARN peuvent **inactiver de grandes régions génomiques** voir des chromosomes entiers en faisant intervenir des modulateurs épigénétiques. (Exemple du lncARN *XIST* détaillé précédemment).

b. Ils peuvent jouer le rôle d'**enhancer** en induisant la transcription en *cis* ou en *trans.* Par exemple, le elncARN Evf2 transcrit dans la région enhancer de ses cibles, interagit avec le facteur de transcription DLX (Distal-Less Homeobox) et la protéine de liaison méthyl-CpG 2 (MECP2) pour réguler de manière épigénétique en *trans* l'expression des gènes Dlx5 et Dlx6 dans le développement des interneurones GABAergiques (Bond et al., 2009).

c. Ils peuvent agir comme un **leurre** aux protéines régulatrices, telles que les facteurs de transcription et les modificateurs de la chromatine, bloquant ainsi leur liaison à l'ADN. Par exemple, le lncARN *MALAT1* (Metastasis-Associated Lung Adenocarcinoma Transcript 1) est fortement exprimé dans les neurones et régule la formation des synapses en modulant l'expression des gènes impliqués dans la formation et / ou le maintien des synapses (Bernard et al., 2010). MALAT1 interagit avec les facteurs d'épissage SR (protéines riches en serine et arginine) et module leur phosphorylation et leur liaison à l'ADN, modifiant le schéma d'épissage d'un sous-ensemble de pré-ARNm endogènes (Tripathi et al., 2010).

d. Ils peuvent servir de **signaux moléculaires** pour activer ou réduire l'expression des gènes en signalant des voies de régulation. Pour le cas de Xist, la présence du lncARN transcrit sur la surface du chromosome X déclenche l'inhibition de la transcription des gènes présents sur ce chromosome suite au recrutement des complexes RNP répressifs ayant une activité méthyl transférase sur les histones (marque H3K27me3)(Brown et al., 1991).

e. Ils peuvent **guider** des protéines, en général des modulateurs de la chromatine vers des sites spécifiques. (Exemple du lncARN *HOTAIR détaillé précédemment).*

f. Ils peuvent agir comme des **échafaudages moléculaires**, liant différentes protéines et formant des complexes ribonucléoprotéiques (RNP), qui affectent également l'expression des gènes. (Exemple du lncARN *HOTAIR* détaillé précédemment*).*

g. Ils **interagissent avec des enzymes**, telles que des kinases, régulant ainsi leur activité catalytique et modifiant leur signalisation. Par exemple, à la suite d'un stress énergétique, le lncARN *NBR2* (*neighbor of BRCA1 gene 2*) interagit avec l'AMPK (AMP-activated protein kinase) et favorise son activité, augmentant ainsi la fonction anti-tumorale de l'AMPK (Liu et al., 2016).

h. Ils peuvent **moduler l'épissage alternatif** d'autres transcrits. Par exemple, *MALAT1* décrit précédemment (c) comme un leurre, qui par son action sur les facteurs d'épissage, modifie l'épissage de certains pré-ARNm (Tripathi et al., 2010). Un autre exemple, le lncARN *MIAT* (Myocardial Infarction Associated Transcript ou *GOMAFU*) présente une séquence répétée en tandem (UACUAAC) similaire à la séquence BPS (intron branch point séquence) des ARNm permettant de lier le facteur d'épissage SF1 (Splicing Factor 1) (Ishii et al., 2006). Ainsi, *MIAT* se fixe au facteur d'épissage SF1 avec une plus grande affinité ce qui affecte la cinétique de la réaction d'épissage des ARNm en modifiant la concentration locale des facteurs d'épissage dans le noyau (Tsuiji et al., 2011).

i. Ils servent d'**éponge à miRNA** empêchant ainsi leur effet de répresseur traductionnel et conduisant à l'augmentation du taux de la protéine cible. Il existe une catégorie de lncARN séquestreurs de

miRNA : les lncARN circulaires (circRNA) dont les extrémités 5' et 3' sont liées de façon covalentes ou les ARN endogènes concurrents (ceARN). Par exemple, le ceARN linc-MD1 (Muscle Differentiation 1) spécifique du muscle, agit comme une éponge pour deux miRNA différents, miR-133 et miR-135. Ces miRNA régulent respectivement l'expression des facteurs de transcription qui activent l'expression de gènes spécifiques du muscle. De cette manière, Linc-MD1 contrôle le moment de la différenciation des myoblastes (Cesana et al., 2011). Un autre exemple est le circRNA, CDR1as (ou ciRS-7, Cerebellar degeneration-related autoantigen 1 antisens) qui contient plus de 70 sites de liaison pour le miR-7 et agit comme une éponge de ce dernier, entrainant une augmentation du taux d'expression de la cible de miR-7 (Hansen et al., 2013).

j. Ils peuvent agir sur l'activité d'une protéine ou modifier sa localisation cellulaire en formant des complexes moléculaires avec celles-ci. Voir l'exemple (g)

k. En ciblant des ARNm, ils peuvent inhiber ou au contraire augmenter leur traduction par les ribosomes. C'est le cas de l'ARNm Zeb2 (Zinc finger E-box-binding homebox 2) dont la traduction peut avoir lieu uniquement en présence du lncARN Zeb2NAT (Zeb2 Natural Antisens Transcript) transcrit en anti-sens du gène Zeb2 (Beltran et al., 2008). Zeb2NAT chevauche le site de reconnaissance de l'épissage de l'intron du pré-ARNm qui contient une séquence IRES (Internal Ribosome Entry Site). Cette séquence IRES permet l'initiation de la traduction. Ainsi, en évitant l'épissage de cette séquence, Zeb2 peut être traduit en protéine et son taux considérablement augmenté. Un autre exemple de lncARN agissant sur la traduction : il existe deux lncARN, BC1 (brain cytoplasmic RNA 1) et BC200 (200 nt brain cytoplasmic RNA), qui peuvent se lier à eIF4A (eukaryotic translation initiation factor 4A), PABP (poly(A)- binding protein) et d'autres facteurs pour bloquer l'assemblage du complexe d'initiation de la traduction requis (Lin et al., 2008; Muddashetty et al., 2002).

l. Certains lncRNA peuvent être transférés à d'autres cellules par des vésicules extracellulaires pour agir sur des molécules à l'intérieure de ses vésicules ou dans d'autres cellules voisines. Des lncARN exosomiques comme Exo2 ou Exo4 sont préférentiellement empaquetés dans des vésicules extracellulaires et peuvent moduler de manière fonctionnelle la viabilité cellulaire par des interactions directes avec la L-lactate déshydrogénase B dans des cellules infectées par le papillomavirus (Hewson et al., 2016). Mais ces lncRNA extracellulaires restent largement inconnus, et la nécessité de davantage de recherches pour caractériser le contenu des vésicules et leurs effets biologiques devient évidente.

m. Ils peuvent être précurseurs de miRNA et être clivés par des RNases. Par exemple, le lncARN *H19* impliqué dans l'empreinte génétique est aussi le précurseur du miR-675 (Bartolomei et al., 1991). Il peut inhiber l'activation et la prolifération des cellules souches hématopoïétiques, en servant de précurseur du miR-675, qui cible le récepteur Igf1r (insulin-like growth factor 1 receptor) dans le placenta (Keniry et al., 2012).

En conclusion, de par leur structure secondaire, les lncARN peuvent se lier à l'ADN, l'ARN, aux miRNA et aux protéines, et agir sur la structure ou les propriétés de ces molécules. Les lncARN se retrouvent impliqués dans la totalité des mécanismes de régulation, ce qui fait d'eux des régulateurs génomiques puissants. L'interrogation fonctionnelle du nombre croissant de nouvelles séquences de lncARN annotées est un défi de taille pour la recherche ces prochaines années. La perturbation de l'expression

des lncARN cibles soit par sur-expression, soit par sous-expression (knockdown=KD) est une approche courante pour étudier leur rôle biologique dans la cellule.

6 Implication des lncARN dans les pathologies humaines

Compte tenu du large éventail de fonctions que révèlent les lncARN dans tous les processus biologiques, il est logique qu'ils soient impliqués dans des pathologies. Des études d'association à l'échelle du génome ont permis d'associer des SNP comme facteurs de risques dans l'étiopathologie de la quasi-totalité des grandes pathologies humaines, comme la maladie d'Alzheimer (MA) (Lambert et al., 2013) ou encore la MP (Nalls et al., 2014). De plus, ces études d'association génomique à grande échelle ont révélé que 88% des SNP associés aux maladies humaines se trouvent en dehors des gènes codants des protéines (Brodie et al., 2016; Hindorff et al., 2009). Outres les études GWAS, des analyses ciblées de l'expression de lncARN d'intérêt par RT-PCRq ou de puces à ADN ont permis d'identifier des lncARN impliqués dans divers processus comme la neurogenèse ou la différenciation neuronale (Ng et al., 2012), mais aussi dans des pathologies. Une analyse des niveaux d'expression de 90 lncARN a mis en évidence une surexpression de différents lncARN, dont le lncARN SNHG1 (Small nucleolar RNA host gene 1), dans les cerveaux de 20 patients souffrant de la MP en comparaison avec 10 patients témoins (Kraus et al., 2017). Puis afin de mieux élucider le rôle de ces lncARN candidats dans les mécanismes qui sous-tendent la maladie, des études fonctionnelles sont misent en place. La plupart des études sur l'implication des lncARN dans les pathologies humaines portent sur les cancers (Brunner et al., 2012; Gibb et al., 2011), mais nous nous intéresserons dans cette partie, à l'implication des lncARN dans la physiopathologie de certaines maladies neurodégénératives en citant quelques exemples :

- La maladie d'Alzheimer (MA) :

Les principales caractéristiques neuro-pathologiques observées dans la MA sont une atrophie du cerveau et une accumulation de plaques amyloïdes extracellulaires et de neurofibrilles intracellulaires enchevêtrées, respectivement composées de protéines *β*-amyloïdes et de protéines TAU (tubulin-associated unit) hyperphosphorylées. La protéine *β*-amyloïde est issue du clivage de la protéine précurseur de l'amyloïde (APP) par l'enzyme BACE1 (*β* site APP-Cleaving Enzyme 1) et la γ-sécrétase initiant une « cascade amyloïde ». Un dysfonctionnement de cette cascade serait à l'origine de la dégénérescence neuronale (Masters et al., 2015). Depuis quelques années, de plus en plus d'études démontrent le rôle de lncARN dans les mécanismes moléculaires associés à la MA.

Par exemple, le lncARN BACE1-AS : l'expression de BACE1 est normalement régulée par le miR-485-5p qui se fixe sur le cadre de lecture ouvert de son pré-ARNm (Figure 15). Or, dans le cortex pariétal et dans le cervelet des patients atteints de la MA, un lncARN BACE1-AS est surexprimé par rapport à des personnes saines (Faghihi et al., 2008). Sur le génome, le gène BACE1-AS chevauche en anti-sens le gène BACE1. Son transcrit peut se lier au pré-ARNm de BACE1 sur le même site de fixation que le miR-485-5p. Ainsi lorsque BACE1-AS se lie au pré-ARNm BACE1, il bloque la fixation du miR-485-5p empêchant la répression de BACE1, ce qui entraine une augmentation de BACE1 et un dérèglement de la cascade amyloïde.

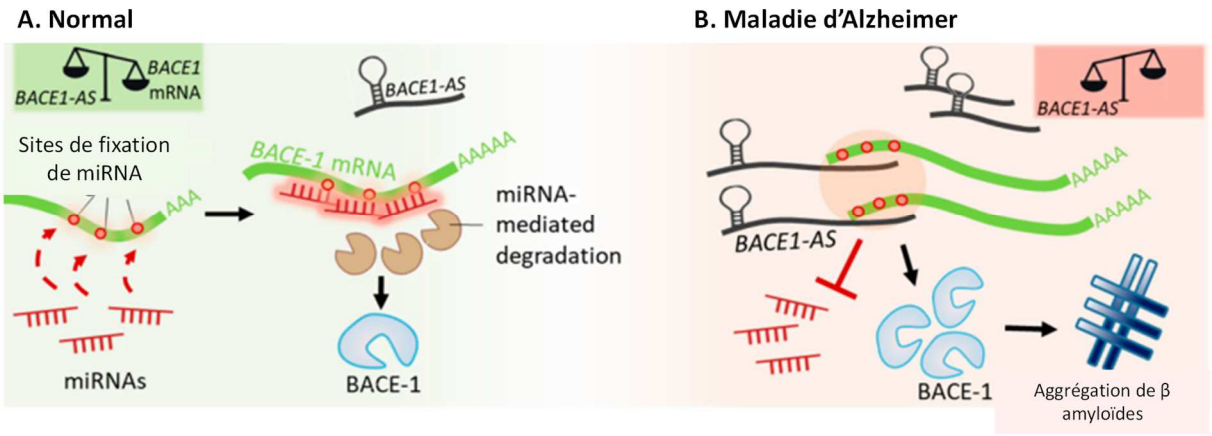


Figure 15 : Représentation schématique du mécanisme d'action du lncARN BACE1-AS (adapté de Pereira Fernandes, 2018). (A) En condition physiologique, l'ARNm de BACE1 est régulé par la fixation de miRNA sur des sites spécifiques de liaison. (B) Dans la maladie d'Alzheimer, le lncARN BACE1-AS empêche la dégradation de BACE1 en se fixant sur les sites de liaisons des miRNA entrainant une augmentation de l'expression de BACE1 et provoquant la formation de plaques amyloïdes.

- La maladie de Huntington (MH) :

Il s'agit d'une pathologie génétique rare et héréditaire, causée par la répétition d'un codon CAG dans le gène codant la protéine huntingtine (HTT). La huntingtine régule la translocation nucléaire du répresseur transcriptionnel REST (RE1-silencing transcription factor). La huntingtine mutée stimule anormalement le trafic nucléaire et cytoplasmique de REST, puis conduit à une dérégulation de l'expression de ses gènes cibles (Zuccato et al., 2003). L'expression du lncARN HTTAS-v1, dont le gène est situé en anti-sens du gène HTT sur le génome et contient la répétition, est réduite dans le cortex de patients MH (Chung et al., 2011). De plus, dans des systèmes cellulaires, la surexpression de HTTAS-v1 réduit spécifiquement l'expression de HTT endogène, alors que sa dégradation par des siRNA (small interfering RNA) augmente les niveaux de transcrits de HTT. Ces travaux suggèrent le rôle régulateur du lncARN HTTAS-v1 sur les niveaux d'expression d'HTT et sa possible implication dans la MH.

- La maladie de Parkinson (MP) :

La MP a longtemps été considérée comme une pathologie sporadique, bien que 10% environ des cas de la MP correspondent à des formes familiales monogéniques. L'une d'elles implique le gène *UCHL1* sur le locus *PARK5* qui code la protéine UCHL1 (Ubiquitin C-Terminal Hydrolase L1), très abondante dans le cerveau et impliquée dans les processus de dégradation des protéines via la voie de l'ubiquitine (Osaka et al., 2003). Sur le génome, ce gène est chevauché en anti-sens par le gène du lncARN *UCHL1-AS* qui possède une séquence SINEUPs (SINE B2 sequence to UP-regulate translation) (Carrieri et al., 2012). Les éléments SINE sont des séquences génomiques répétées et l'élément B2 permet au transcrit qui le porte de cibler spécifiquement l'ARNm qu'il chevauche pour augmenter sa traduction. Ainsi, en condition de stress, le lncARN *UCHL1-AS* physiologiquement majoritaire dans le noyau des neurones DA migre dans le cytoplasme et augmente la traduction de la protéine UCHL1 en ciblant son transcrit spécifiquement (Carrieri et al., 2012).

Un autre lncARN, na*PINK1* (natural antisens PTEN Induced Putative Kinase 1) situé sur le locus du gène *PINK1*, stabilise l'expression d'une isoforme de *PINK1* (svPINK1) dans les neurones pendant la biogenèse mitochondriale (Scheele et al., 2007). L'inactivation de *PINK1* entraîne un changement de morphologie des mitochondries et la libération de dopamine, détruisant les cellules nerveuses dopaminergiques

(Morais et al., 2009). La protéine PINK1 a aussi un rôle protecteur suite à un stress mitochondrial, en guidant les mitochondries défectueuses vers leur dégradation par mitophagie (Szargel et al., 2016). Dans un contexte de MP, l'autophagie induite par le MPTP (voir partie Facteurs environnementaux), est favorisée par le lncARN *NEAT1* (Nuclear Paraspeckle Assembly Transcript 1) grâce à la stabilisation de la protéine PINK1 (Yan et al., 2018).

Récemment, il a été montré que l'expression élevée du lncARN *HOTAIR* (Hox transcript antisense intergenic RNA) favorise l'apparition de la MP induite par le MPTP dans un modèle de souris en stabilisant l'expression de *LRRK2* (Leucine Rich Repeat Kinase 2) dont la mutation est associée à la MP (Wang et al., 2017). De plus, l'inhibition de *HOTAIR* dans ce modèle, protège contre l'apoptose neuronale en réprimant l'activité de la caspase 3. La découverte que HOTAIR favorise la MP induite par le MPTP pourrait aider à la compréhension des mécanismes moléculaires impliqués dans la MP.

En conclusion, l'accroissement du nombre d'études de lncARN dérégulés dans les pathologies neurodégénératives humaines, va permettre une meilleure compréhension de leurs mécanismes d'action et de régulation. Ces connaissances vont aider à développer de nouveaux diagnostics et offrir de nouvelles cibles thérapeutiques utiles.

Références bibliographiques

Agarwal, A., Koppstein, D., Rozowsky, J., Sboner, A., Habegger, L., Hillier, L.W., Sasidharan, R., Reinke, V., Waterston, R.H., and Gerstein, M. (2010). Comparison and calibration of transcriptome data from RNA-Seq and tiling arrays. BMC Genomics *11*, 383.

Ayupe, A.C., Tahira, A.C., Camargo, L., Beckedorff, F.C., Verjovski-Almeida, S., and Reis, E.M. (2015). Global analysis of biogenesis, stability and sub-cellular localization of lncRNAs mapping to intragenic regions of the human genome. RNA Biol *12*, 877-892.

Banfai, B., Jia, H., Khatun, J., Wood, E., Risk, B., Gundling, W.E., Jr., Kundaje, A., Gunawardena, H.P., Yu, Y., Xie, L.*, et al.* (2012). Long noncoding RNAs are rarely translated in two human cell lines. Genome Res *22*, 1646-1657.

Barker, R.A., Studer, L., Cattaneo, E., Takahashi, J., and consortium, G.F.P. (2015). G-Force PD: a global initiative in coordinating stem cell-based dopamine treatments for Parkinson's disease. NPJ Parkinsons Dis *1*, 15017.

Bartolomei, M.S., Zemel, S., and Tilghman, S.M. (1991). Parental imprinting of the mouse H19 gene. Nature *351*, 153-155.

Bartus, R.T., and Johnson, E.M., Jr. (2017a). Clinical tests of neurotrophic factors for human neurodegenerative diseases, part 1: Where have we been and what have we learned? Neurobiol Dis *97*, 156-168.

Bartus, R.T., and Johnson, E.M., Jr. (2017b). Clinical tests of neurotrophic factors for human neurodegenerative diseases, part 2: Where do we stand and where must we go next? Neurobiol Dis *97*, 169-178.

Beltran, M., Puig, I., Pena, C., Garcia, J.M., Alvarez, A.B., Pena, R., Bonilla, F., and de Herreros, A.G. (2008). A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snail1-induced epithelial-mesenchymal transition. Genes Dev *22*, 756-769.

Berg, D., Postuma, R.B., Adler, C.H., Bloem, B.R., Chan, P., Dubois, B., Gasser, T., Goetz, C.G., Halliday, G., Joseph, L.*, et al.* (2015). MDS research criteria for prodromal Parkinson's disease. Mov Disord *30*, 1600-1611.

Bernard, D., Prasanth, K.V., Tripathi, V., Colasse, S., Nakamura, T., Xuan, Z., Zhang, M.Q., Sedel, F., Jourdren, L., Coulpier, F.*, et al.* (2010). A long nuclear-retained non-coding RNA regulates synaptogenesis by modulating gene expression. EMBO J *29*, 3082-3093.

Bertone, P., Stolc, V., Royce, T.E., Rozowsky, J.S., Urban, A.E., Zhu, X., Rinn, J.L., Tongprasit, W., Samanta, M., Weissman, S.*, et al.* (2004). Global identification of human transcribed sequences with genome tiling arrays. Science *306*, 2242-2246.

Betarbet, R., Sherer, T.B., MacKenzie, G., Garcia-Osuna, M., Panov, A.V., and Greenamyre, J.T. (2000). Chronic systemic pesticide exposure reproduces features of Parkinson's disease. Nat Neurosci *3*, 1301-1306.

Beyer, M.K., Herlofson, K., Arsland, D., and Larsen, J.P. (2001). Causes of death in a community-based study of Parkinson's disease. Acta Neurol Scand *103*, 7-11.

Bond, A.M., Vangompel, M.J., Sametsky, E.A., Clark, M.F., Savage, J.C., Disterhoft, J.F., and Kohtz, J.D. (2009). Balanced gene regulation by an embryonic brain ncRNA is critical for adult hippocampal GABA circuitry. Nat Neurosci *12*, 1020-1027.

Bower, J.H., Maraganore, D.M., Peterson, B.J., McDonnell, S.K., Ahlskog, J.E., and Rocca, W.A. (2003). Head trauma preceding PD: a case-control study. Neurology *60*, 1610-1615.

Braak, H., Del Tredici, K., Rub, U., de Vos, R.A., Jansen Steur, E.N., and Braak, E. (2003). Staging of brain pathology related to sporadic Parkinson's disease. Neurobiol Aging *24*, 197-211.

Brizard, M., Carcenac, C., Bemelmans, A.P., Feuerstein, C., Mallet, J., and Savasta, M. (2006). Functional reinnervation from remaining DA terminals induced by GDNF lentivirus in a rat model of early Parkinson's disease. Neurobiol Dis *21*, 90-101.

Brodie, A., Azaria, J.R., and Ofran, Y. (2016). How far from the SNP may the causative genes be? Nucleic Acids Res *44*, 6046-6054.

Bronstein, J.M., Tagliati, M., Alterman, R.L., Lozano, A.M., Volkmann, J., Stefani, A., Horak, F.B., Okun, M.S., Foote, K.D., Krack, P.*, et al.* (2011). Deep brain stimulation for Parkinson disease: an expert consensus and review of key issues. Arch Neurol *68*, 165.

Brown, C.J., Ballabio, A., Rupert, J.L., Lafreniere, R.G., Grompe, M., Tonlorenzi, R., and Willard, H.F. (1991). A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. Nature *349*, 38-44.

Brunner, A.L., Beck, A.H., Edris, B., Sweeney, R.T., Zhu, S.X., Li, R., Montgomery, K., Varma, S., Gilks, T., Guo, X.*, et al.* (2012). Transcriptional profiling of long non-coding RNAs and novel transcribed regions across a diverse panel of archived human cancers. Genome Biol *13*, R75.

Burre, J. (2015). The Synaptic Function of alpha-Synuclein. J Parkinsons Dis *5*, 699-713.

Cabili, M.N., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., Regev, A., and Rinn, J.L. (2011). Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev *25*, 1915-1927.

Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C.*, et al.* (2005). The transcriptional landscape of the mammalian genome. Science *309*, 1559-1563.
Carrieri, C., Cimatti, L., Biagioli, M., Beugnet, A., Zucchelli, S., Fedele, S., Pesce, E., Ferrer, I., Collavin, L., Santoro, C.*, et al.* (2012). Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. Nature *491*, 454-457.
Castel, S.E., and Martienssen, R.A. (2013). RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. Nat Rev Genet *14*, 100-112.
Cenci, M.A. (2014). Presynaptic Mechanisms of l-DOPA-Induced Dyskinesia: The Findings, the Debate, and the Therapeutic Implications. Front Neurol *5*, 242.
Cesana, M., Cacchiarelli, D., Legnini, I., Santini, T., Sthandier, O., Chinappi, M., Tramontano, A., and Bozzoni, I. (2011). A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. Cell *147*, 358-369.
Chung, D.W., Rudnicki, D.D., Yu, L., and Margolis, R.L. (2011). A natural antisense transcript at the Huntington's disease repeat locus regulates HTT expression. Hum Mol Genet *20*, 3467-3477.
Clemson, C.M., Hutchinson, J.N., Sara, S.A., Ensminger, A.W., Fox, A.H., Chess, A., and Lawrence, J.B. (2009). An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. Mol Cell *33*, 717-726.
Connolly, B., and Fox, S.H. (2014). Treatment of cognitive, psychiatric, and affective disorders associated with Parkinson's disease. Neurotherapeutics *11*, 78-91.
Consortium, E.P. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57-74.
Corti, O., Lesage, S., and Brice, A. (2011). What genetics tells us about the causes and mechanisms of Parkinson's disease. Physiol Rev *91*, 1161-1218.
Coupland, K.G., Kim, W.S., Halliday, G.M., Hallupp, M., Dobson-Stone, C., and Kwok, J.B. (2016). Role of the Long Non-Coding RNA MAPT-AS1 in Regulation of Microtubule Associated Protein Tau (MAPT) Expression in Parkinson's Disease. PLoS One *11*, e0157924.
de Rijk, M.C., Rocca, W.A., Anderson, D.W., Melcon, M.O., Breteler, M.M., and Maraganore, D.M. (1997). A population perspective on diagnostic criteria for Parkinson's disease. Neurology *48*, 1277-1281.
Deng, H., Wang, P., and Jankovic, J. (2018). The genetics of Parkinson disease. Ageing Res Rev *42*, 72-85.
Derrien, T., Guigo, R., and Johnson, R. (2011). The Long Non-Coding RNAs: A New (P)layer in the "Dark Matter". Front Genet *2*, 107.
Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., Guernec, G., Martin, D., Merkel, A., Knowles, D.G.*, et al.* (2012). The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res *22*, 1775-1789.
Deuschl, G., and Agid, Y. (2013). Subthalamic neurostimulation for Parkinson's disease with early fluctuations: balancing the risks and benefits. Lancet Neurol *12*, 1025-1034.
Dorsey, E.R., Constantinescu, R., Thompson, J.P., Biglan, K.M., Holloway, R.G., Kieburtz, K., Marshall, F.J., Ravina, B.M., Schifitto, G., Siderowf, A.*, et al.* (2007). Projected number of people with Parkinson disease in the most populous nations, 2005 through 2030. Neurology *68*, 384-386.
Elbaz, A., Carcaillon, L., Kab, S., and Moisan, F. (2016). Epidemiology of Parkinson's disease. Rev Neurol (Paris) *172*, 14-26.
Faghihi, M.A., Modarresi, F., Khalil, A.M., Wood, D.E., Sahagan, B.G., Morgan, T.E., Finch, C.E., St Laurent, G., 3rd, Kenny, P.J., and 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*, 723-730.
Fang, F., Chen, H., Feldman, A.L., Kamel, F., Ye, W., and Wirdefeldt, K. (2012). Head injury and Parkinson's disease: a population-based study. Mov Disord *27*, 1632-1635.
Fox, S.H., Katzenschlager, R., Lim, S.Y., Ravina, B., Seppi, K., Coelho, M., Poewe, W., Rascol, O., Goetz, C.G., and Sampaio, C. (2011). The Movement Disorder Society Evidence-Based Medicine Review Update: Treatments for the motor symptoms of Parkinson's disease. Mov Disord *26 Suppl 3*, S2-41.
Gash, D.M., Rutland, K., Hudson, N.L., Sullivan, P.G., Bing, G., Cass, W.A., Pandya, J.D., Liu, M., Choi, D.Y., Hunter, R.L.*, et al.* (2008). Trichloroethylene: Parkinsonism and complex 1 mitochondrial neurotoxicity. Ann Neurol *63*, 184-192.
Geisler, S., and Coller, J. (2013). RNA in unexpected places: long non-coding RNA functions in diverse cellular contexts. Nat Rev Mol Cell Biol *14*, 699-712.
Giannakakis, A., Zhang, J., Jenjaroenpun, P., Nama, S., Zainolabidin, N., Aau, M.Y., Yarmishyn, A.A., Vaz, C., Ivshina, A.V., Grinchuk, O.V.*, et al.* (2015). Contrasting expression patterns of coding and noncoding parts of the human genome upon oxidative stress. Sci Rep *5*, 9737.

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., *et al.* (2011). Human cancer long non-coding RNA transcriptomes. PLoS One *6*, e25915.
Goedert, M., Spillantini, M.G., Del Tredici, K., and Braak, H. (2013). 100 years of Lewy pathology. Nat Rev Neurol *9*, 13-24.
Goldman, S.M., Kamel, F., Ross, G.W., Jewell, S.A., Bhudhikanok, G.S., Umbach, D., Marras, C., Hauser, R.A., Jankovic, J., Factor, S.A., *et al.* (2012). Head injury, alpha-synuclein Rep1, and Parkinson's disease. Ann Neurol *71*, 40-48.
Goldman, S.M., Tanner, C.M., Oakes, D., Bhudhikanok, G.S., Gupta, A., and Langston, J.W. (2006). Head injury and Parkinson's disease risk in twins. Ann Neurol *60*, 65-72.
Guenzl, P.M., and Barlow, D.P. (2012). Macro lncRNAs: a new layer of cis-regulatory information in the mammalian genome. RNA Biol *9*, 731-741.
Guttman, M., Amit, I., Garber, M., French, C., Lin, M.F., Feldser, D., Huarte, M., Zuk, O., Carey, B.W., Cassady, J.P., *et al.* (2009). Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature *458*, 223-227.
Guttman, M., and Rinn, J.L. (2012). Modular regulatory principles of large non-coding RNAs. Nature *482*, 339-346.
Guttman, M., Russell, P., Ingolia, N.T., Weissman, J.S., and Lander, E.S. (2013). Ribosome profiling provides evidence that large noncoding RNAs do not encode proteins. Cell *154*, 240-251.
Hansen, T.B., Jensen, T.I., Clausen, B.H., Bramsen, J.B., Finsen, B., Damgaard, C.K., and Kjems, J. (2013). Natural RNA circles function as efficient microRNA sponges. Nature *495*, 384-388.
Harris, M.A., Shen, H., Marion, S.A., Tsui, J.K., and Teschke, K. (2013). Head injuries and Parkinson's disease in a case-control study. Occup Environ Med *70*, 839-844.
Hewson, C., Capraro, D., Burdach, J., Whitaker, N., and Morris, K.V. (2016). Extracellular vesicle associated long non-coding RNAs functionally enhance cell viability. Noncoding RNA Res *1*, 3-11.
Hezroni, H., Koppstein, D., Schwartz, M.G., Avrutin, A., Bartel, D.P., and Ulitsky, I. (2015). Principles of long noncoding RNA evolution derived from direct comparison of transcriptomes in 17 species. Cell Rep *11*, 1110-1122.
Hindorff, L.A., Sethupathy, P., Junkins, H.A., Ramos, E.M., Mehta, J.P., Collins, F.S., and Manolio, T.A. (2009). Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc Natl Acad Sci U S A *106*, 9362-9367.
Housman, G., and Ulitsky, I. (2016). Methods for distinguishing between protein-coding and long noncoding RNAs and the elusive biological purpose of translation of long noncoding RNAs. Biochim Biophys Acta *1859*, 31-40.
Ingolia, N.T. (2014). Ribosome profiling: new views of translation, from single codons to genome scale. Nat Rev Genet *15*, 205-213.
Ishii, N., Ozaki, K., Sato, H., Mizuno, H., Saito, S., Takahashi, A., Miyamoto, Y., Ikegawa, S., Kamatani, N., Hori, M., *et al.* (2006). Identification of a novel non-coding RNA, MIAT, that confers risk of myocardial infarction. J Hum Genet *51*, 1087-1099.
Jakes, R., Spillantini, M.G., and Goedert, M. (1994). Identification of two distinct synucleins from human brain. FEBS Lett *345*, 27-32.
Jarroux, J., Morillon, A., and Pinskaya, M. (2017). History, Discovery, and Classification of lncRNAs. Adv Exp Med Biol *1008*, 1-46.
Kab, S., Spinosi, J., Chaperon, L., Dugravot, A., Singh-Manoux, A., Moisan, F., and Elbaz, A. (2017). Agricultural activities and the incidence of Parkinson's disease in the general French population. Eur J Epidemiol *32*, 203-216.
Kalia, L.V., and Lang, A.E. (2015). Parkinson's disease. Lancet *386*, 896-912.
Kalia, L.V., and Lang, A.E. (2016). Parkinson disease in 2015: Evolving basic, pathological and clinical concepts in PD. Nat Rev Neurol *12*, 65-66.
Kalinderi, K., Bostantjopoulou, S., and Fidani, L. (2016). The genetic background of Parkinson's disease: current progress and future prospects. Acta Neurol Scand *134*, 314-326.
Kapranov, P., Cheng, J., Dike, S., Nix, D.A., Duttagupta, R., Willingham, A.T., Stadler, P.F., Hertel, J., Hackermuller, J., Hofacker, I.L., *et al.* (2007). RNA maps reveal new RNA classes and a possible function for pervasive transcription. Science *316*, 1484-1488.
Keniry, A., Oxley, D., Monnier, P., Kyba, M., Dandolo, L., Smits, G., and Reik, W. (2012). The H19 lincRNA is a developmental reservoir of miR-675 that suppresses growth and Igf1r. Nat Cell Biol *14*, 659-665.

Khalil, A.M., Guttman, M., Huarte, M., Garber, M., Raj, A., Rivea Morales, D., Thomas, K., Presser, A., Bernstein, B.E., van Oudenaarden, A*., et al.* (2009). Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. Proc Natl Acad Sci U S A *106*, 11667-11672.
Kirkeby, A., Grealish, S., Wolf, D.A., Nelander, J., Wood, J., Lundblad, M., Lindvall, O., and Parmar, M. (2012). Generation of regionally specified neural progenitors and functional neurons from human embryonic stem cells under defined conditions. Cell Rep *1*, 703-714.
Kitada, T., Asakawa, S., Hattori, N., Matsumine, H., Yamamura, Y., Minoshima, S., Yokochi, M., Mizuno, Y., and Shimizu, N. (1998). Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. Nature *392*, 605-608.
Konno, T., Ross, O.A., Puschmann, A., Dickson, D.W., and Wszolek, Z.K. (2016). Autosomal dominant Parkinson's disease caused by SNCA duplications. Parkinsonism Relat Disord *22 Suppl 1*, S1-6.
Kordower, J.H., and Bjorklund, A. (2013). Trophic factor gene therapy for Parkinson's disease. Mov Disord *28*, 96-109.
Kraus, T.F.J., Haider, M., Spanner, J., Steinmaurer, M., Dietinger, V., and Kretzschmar, H.A. (2017). Altered Long Noncoding RNA Expression Precedes the Course of Parkinson's Disease-a Preliminary Report. Mol Neurobiol *54*, 2869-2877.
Kriks, S., Shim, J.W., Piao, J., Ganat, Y.M., Wakeman, D.R., Xie, Z., Carrillo-Reid, L., Auyeung, G., Antonacci, C., Buch, A*., et al.* (2011). Dopamine neurons derived from human ES cells efficiently engraft in animal models of Parkinson's disease. Nature *480*, 547-551.
Kuhn, A.A., and Volkmann, J. (2017). Innovations in deep brain stimulation methodology. Mov Disord *32*, 11-19.
Lambert, J.C., Ibrahim-Verbaas, C.A., Harold, D., Naj, A.C., Sims, R., Bellenguez, C., DeStafano, A.L., Bis, J.C., Beecham, G.W., Grenier-Boley, B*., et al.* (2013). Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet *45*, 1452-1458.
Langston, J.W., and Ballard, P.A., Jr. (1983). Parkinson's disease in a chemist working with 1-methyl-4-phenyl-1,2,5,6-tetrahydropyridine. N Engl J Med *309*, 310.
Lazarou, M., Sliter, D.A., Kane, L.A., Sarraf, S.A., Wang, C., Burman, J.L., Sideris, D.P., Fogel, A.I., and Youle, R.J. (2015). The ubiquitin kinase PINK1 recruits autophagy receptors to induce mitophagy. Nature *524*, 309-314.
Lesage, S., and Brice, A. (2009). Parkinson's disease: from monogenic forms to genetic susceptibility factors. Hum Mol Genet *18*, R48-59.
LeWitt, P.A., and Fahn, S. (2016). Levodopa therapy for Parkinson disease: A look backward and forward. Neurology *86*, S3-12.
Lin, D., Pestova, T.V., Hellen, C.U., and Tiedge, H. (2008). Translational control by a small RNA: dendritic BC1 RNA targets the eukaryotic initiation factor 4A helicase mechanism. Mol Cell Biol *28*, 3008-3019.
Lionnet, A., Leclair-Visonneau, L., Neunlist, M., Murayama, S., Takao, M., Adler, C.H., Derkinderen, P., and Beach, T.G. (2018). Does Parkinson's disease start in the gut? Acta Neuropathol *135*, 1-12.
Liu, X., Xiao, Z.D., and Gan, B. (2016). An lncRNA switch for AMPK activation. Cell Cycle *15*, 1948-1949.
Maiti, P., Manna, J., and Dunbar, G.L. (2017). Current understanding of the molecular mechanisms in Parkinson's disease: Targets for potential treatments. Transl Neurodegener *6*, 28.
Man, J.H.K., Groenink, L., and Caiazzo, M. (2018). Cell reprogramming approaches in gene- and cell-based therapies for Parkinson's disease. J Control Release *286*, 114-124.
Mandler, M., Valera, E., Rockenstein, E., Weninger, H., Patrick, C., Adame, A., Santic, R., Meindl, S., Vigl, B., Smrzka, O*., et al.* (2014). Next-generation active immunization approach for synucleinopathies: implications for Parkinson's disease clinical trials. Acta Neuropathol *127*, 861-879.
Manning-Bog, A.B., McCormack, A.L., Li, J., Uversky, V.N., Fink, A.L., and Di Monte, D.A. (2002). The herbicide paraquat causes up-regulation and aggregation of alpha-synuclein in mice: paraquat and alpha-synuclein. J Biol Chem *277*, 1641-1644.
Maroteaux, L., Campanelli, J.T., and Scheller, R.H. (1988). Synuclein: a neuron-specific protein localized to the nucleus and presynaptic nerve terminal. J Neurosci *8*, 2804-2815.
Masters, C.L., Bateman, R., Blennow, K., Rowe, C.C., Sperling, R.A., and Cummings, J.L. (2015). Alzheimer's disease. Nat Rev Dis Primers *1*, 15056.
Matsumine, H., Saito, M., Shimoda-Matsubayashi, S., Tanaka, H., Ishikawa, A., Nakagawa-Hattori, Y., Yokochi, M., Kobayashi, T., Igarashi, S., Takano, H*., et al.* (1997). Localization of a gene for an autosomal recessive form of juvenile Parkinsonism to chromosome 6q25.2-27. Am J Hum Genet *60*, 588-596.
McGuire Kuhl, M. (2016) Vaccine for Parkinson's reports positive results from boost study. MichaelJFox.org
Mele, M., Mattioli, K., Mallard, W., Shechner, D.M., Gerhardinger, C., and Rinn, J.L. (2017). Chromatin environment, transcriptional regulation, and splicing distinguish lincRNAs and mRNAs. Genome Res *27*, 27-37.

Mercer, T.R., Gerhardt, D.J., Dinger, M.E., Crawford, J., Trapnell, C., Jeddeloh, J.A., Mattick, J.S., and Rinn, J.L. (2011). Targeted RNA sequencing reveals the deep complexity of the human transcriptome. Nat Biotechnol *30*, 99-104.

Moisan, F., Gourlet, V., Mazurie, J.L., Dupupet, J.L., Houssinot, J., Goldberg, M., Imbernon, E., Tzourio, C., and Elbaz, A. (2011). Prediction model of Parkinson's disease based on antiparkinsonian drug claims. Am J Epidemiol *174*, 354-363.

Moisan F, Kab S, Moutengou E, Boussac-Zerebska M, Carcaillon-Bentata L, Elbaz A. (2018) Fréquence de la maladie de Parkinson en France. Données nationales et régionales 2010-2015. Saint-Maurice : Santé publique France

Morais, V.A., Verstreken, P., Roethig, A., Smet, J., Snellinx, A., Vanbrabant, M., Haddad, D., Frezza, C., Mandemakers, W., Vogt-Weisenhorn, D*., et al.* (2009). Parkinson's disease mutations in PINK1 result in decreased Complex I activity and deficient synaptic function. EMBO Mol Med *1*, 99-111.

Moran, I., Akerman, I., van de Bunt, M., Xie, R., Benazra, M., Nammo, T., Arnes, L., Nakic, N., Garcia-Hurtado, J., Rodriguez-Segui, S*., et al.* (2012). Human beta cell transcriptome analysis uncovers lncRNAs that are tissue-specific, dynamically regulated, and abnormally expressed in type 2 diabetes. Cell Metab *16*, 435-448.

Muddashetty, R., Khanam, T., Kondrashov, A., Bundman, M., Iacoangeli, A., Kremerskothen, J., Duning, K., Barnekow, A., Huttenhofer, A., Tiedge, H*., et al.* (2002). Poly(A)-binding protein is associated with neuronal BC1 and BC200 ribonucleoprotein particles. J Mol Biol *321*, 433-445.

Mullin, S., and Schapira, A. (2015). The genetics of Parkinson's disease. Br Med Bull *114*, 39-52.

Nalls, M.A., Pankratz, N., Lill, C.M., Do, C.B., Hernandez, D.G., Saad, M., DeStefano, A.L., Kara, E., Bras, J., Sharma, M*., et al.* (2014). Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. Nat Genet *46*, 989-993.

Nemani, V.M., Lu, W., Berge, V., Nakamura, K., Onoa, B., Lee, M.K., Chaudhry, F.A., Nicoll, R.A., and Edwards, R.H. (2010). Increased expression of alpha-synuclein reduces neurotransmitter release by inhibiting synaptic vesicle reclustering after endocytosis. Neuron *65*, 66-79.

Ng, S.Y., Johnson, R., and Stanton, L.W. (2012). Human long non-coding RNAs promote pluripotency and neuronal differentiation by association with chromatin modifiers and transcription factors. EMBO J *31*, 522-533.

Noh, J.H., Kim, K.M., Abdelmohsen, K., Yoon, J.H., Panda, A.C., Munk, R., Kim, J., Curtis, J., Moad, C.A., Wohler, C.M*., et al.* (2016). HuR and GRSF1 modulate the nuclear export and mitochondrial localization of the lncRNA RMRP. Genes Dev *30*, 1224-1239.

Nuytemans, K., Theuns, J., Cruts, M., and Van Broeckhoven, C. (2010). Genetic etiology of Parkinson disease associated with mutations in the SNCA, PARK2, PINK1, PARK7, and LRRK2 genes: a mutation update. Hum Mutat *31*, 763-780.

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Osato, N., Saito, R., Suzuki, H*., et al.* (2002). Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. Nature *420*, 563-573.

Osaka, H., Wang, Y.L., Takada, K., Takizawa, S., Setsuie, R., Li, H., Sato, Y., Nishikawa, K., Sun, Y.J., Sakurai, M*., et al.* (2003). Ubiquitin carboxy-terminal hydrolase L1 binds to and stabilizes monoubiquitin in neuron. Hum Mol Genet *12*, 1945-1958.

Parkinson, J. (2002). An essay on the shaking palsy. 1817. J Neuropsychiatry Clin Neurosci *14*, 223-236; discussion 222.

Pauli, A., Valen, E., Lin, M.F., Garber, M., Vastenhouw, N.L., Levin, J.Z., Fan, L., Sandelin, A., Rinn, J.L., Regev, A*., et al.* (2012). Systematic identification of long noncoding RNAs expressed during zebrafish embryogenesis. Genome Res *22*, 577-591.

Pepeu, G., and Grazia Giovannini, M. (2017). The fate of the brain cholinergic neurons in neurodegenerative diseases. Brain Res *1670*, 173-184.

Pezzoli, G., and Cereda, E. (2013). Exposure to pesticides or solvents and risk of Parkinson disease. Neurology *80*, 2035-2041.

Poliseno, L., Salmena, L., Zhang, J., Carver, B., Haveman, W.J., and Pandolfi, P.P. (2010). A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. Nature *465*, 1033-1038.

Polymeropoulos, M.H., Lavedan, C., Leroy, E., Ide, S.E., Dehejia, A., Dutra, A., Pike, B., Root, H., Rubenstein, J., Boyer, R*., et al.* (1997). Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. Science *276*, 2045-2047.

Ponjavic, J., Ponting, C.P., and Lunter, G. (2007). Functionality or transcriptional noise? Evidence for selection within long noncoding RNAs. Genome Res *17*, 556-565.

Porro, A., Feuerhahn, S., Reichenbach, P., and Lingner, J. (2010). Molecular dissection of telomeric repeat-containing RNA biogenesis unveils the presence of distinct and multiple regulatory pathways. Mol Cell Biol *30*, 4808-4817.
Postuma, R.B., Berg, D., Stern, M., Poewe, W., Olanow, C.W., Oertel, W., Obeso, J., Marek, K., Litvan, I., Lang, A.E.*, et al.* (2015). MDS clinical diagnostic criteria for Parkinson's disease. Mov Disord *30*, 1591-1601.
Pringsheim, T., Jette, N., Frolkis, A., and Steeves, T.D. (2014). The prevalence of Parkinson's disease: a systematic review and meta-analysis. Mov Disord *29*, 1583-1590.
Pruitt, K.D., Tatusova, T., Brown, G.R., and Maglott, D.R. (2012). NCBI Reference Sequences (RefSeq): current status, new features and genome annotation policy. Nucleic Acids Res *40*, D130-135.
Rackham, O., Shearwood, A.M., Mercer, T.R., Davies, S.M., Mattick, J.S., and Filipovska, A. (2011). Long noncoding RNAs are generated from the mitochondrial genome and regulated by nuclear-encoded proteins. RNA *17*, 2085-2093.
Ravasi, T., Suzuki, H., Pang, K.C., Katayama, S., Furuno, M., Okunishi, R., Fukuda, S., Ru, K., Frith, M.C., Gongora, M.M.*, et al.* (2006). Experimental validation of the regulated expression of large numbers of non-coding RNAs from the mouse genome. Genome Res *16*, 11-19.
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.*, et al.* (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell *129*, 1311-1323.
Ryskalin, L., Busceti, C.L., Limanaqi, F., Biagioni, F., Gambardella, S., and Fornai, F. (2018). A Focus on the Beneficial Effects of Alpha Synuclein and a Re-Appraisal of Synucleinopathies. Curr Protein Pept Sci *19*, 598-611.
Scheele, C., Petrovic, N., Faghihi, M.A., Lassmann, T., Fredriksson, K., Rooyackers, O., Wahlestedt, C., Good, L., and Timmons, J.A. (2007). The human PINK1 locus is regulated in vivo by a non-coding natural antisense RNA during modulation of mitochondrial function. BMC Genomics *8*, 74.
Shannon, K.M., Keshavarzian, A., Dodiya, H.B., Jakate, S., and Kordower, J.H. (2012). Is alpha-synuclein in the colon a biomarker for premotor Parkinson's disease? Evidence from 3 cases. Mov Disord *27*, 716-719.
Smith, K.N., Starmer, J., Miller, S.C., Sethupathy, P., and Magnuson, T. (2017). Long Noncoding RNA Moderates MicroRNA Activity to Maintain Self-Renewal in Embryonic Stem Cells. Stem Cell Reports *9*, 108-121.
Song, X., Wang, X., Arai, S., and Kurokawa, R. (2012). Promoter-associated noncoding RNA from the CCND1 promoter. Methods Mol Biol *809*, 609-622.
St Laurent, G., Vyatkin, Y., Antonets, D., Ri, M., Qi, Y., Saik, O., Shtokalo, D., de Hoon, M.J., Kawaji, H., Itoh, M.*, et al.* (2016). Functional annotation of the vlinc class of non-coding RNAs using systems biology approach. Nucleic Acids Res *44*, 3233-3252.
St Laurent, G., Wahlestedt, C., and Kapranov, P. (2015). The Landscape of long noncoding RNA classification. Trends Genet *31*, 239-251.
Szargel, R., Shani, V., Abd Elghani, F., Mekies, L.N., Liani, E., Rott, R., and Engelender, S. (2016). The PINK1, synphilin-1 and SIAH-1 complex constitutes a novel mitophagy pathway. Hum Mol Genet *25*, 3476-3490.
Tanner, C.M., Kamel, F., Ross, G.W., Hoppin, J.A., Goldman, S.M., Korell, M., Marras, C., Bhudhikanok, G.S., Kasten, M., Chade, A.R.*, et al.* (2011). Rotenone, paraquat, and Parkinson's disease. Environ Health Perspect *119*, 866-872.
Thannickal, T.C., Lai, Y.Y., and Siegel, J.M. (2007). Hypocretin (orexin) cell loss in Parkinson's disease. Brain *130*, 1586-1595.
Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol *28*, 511-515.
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.*, et al.* (2010). The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. Mol Cell *39*, 925-938.
Tsuiji, H., Yoshimoto, R., Hasegawa, Y., Furuno, M., Yoshida, M., and Nakagawa, S. (2011). Competition between a noncoding exon and introns: Gomafu contains tandem UACUAAC repeats and associates with splicing factor-1. Genes Cells *16*, 479-490.
Ulitsky, I., and Bartel, D.P. (2013). lincRNAs: genomics, evolution, and mechanisms. Cell *154*, 26-46.
Ulitsky, I., Shkumatava, A., Jan, C.H., Sive, H., and Bartel, D.P. (2011). Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. Cell *147*, 1537-1550.
Valente, E.M., Bentivoglio, A.R., Dixon, P.H., Ferraris, A., Ialongo, T., Frontali, M., Albanese, A., and Wood, N.W. (2001). Localization of a novel locus for autosomal recessive early-onset parkinsonism, PARK6, on human chromosome 1p35-p36. Am J Hum Genet *68*, 895-900.

van der Mark, M., Brouwer, M., Kromhout, H., Nijssen, P., Huss, A., and Vermeulen, R. (2012). Is pesticide use related to Parkinson disease? Some clues to heterogeneity in study results. Environ Health Perspect *120*, 340-347.
van der Merwe, C., Jalali Sefid Dashti, Z., Christoffels, A., Loos, B., and Bardien, S. (2015). Evidence for a common biological pathway linking three Parkinson's disease-causing genes: parkin, PINK1 and DJ-1. Eur J Neurosci *41*, 1113-1125.
van Heesch, S., van Iterson, M., Jacobi, J., Boymans, S., Essers, P.B., de Bruijn, E., Hao, W., MacInnes, A.W., Cuppen, E., and Simonis, M. (2014). Extensive localization of long noncoding RNAs to the cytosol and mono- and polyribosomal complexes. Genome Biol *15*, R6.
Van Maele-Fabry, G., Hoet, P., Vilain, F., and Lison, D. (2012). Occupational exposure to pesticides and Parkinson's disease: a systematic review and meta-analysis of cohort studies. Environ Int *46*, 30-43.
van Rossum, D., Verheijen, B.M., and Pasterkamp, R.J. (2016). Circular RNAs: Novel Regulators of Neuronal Development. Front Mol Neurosci *9*, 74.
Vanle, B., Olcott, W., Jimenez, J., Bashmi, L., Danovitch, I., and IsHak, W.W. (2018). NMDA antagonists for treating the non-motor symptoms in Parkinson's disease. Transl Psychiatry *8*, 117.
Verstraeten, A., Theuns, J., and Van Broeckhoven, C. (2015). Progress in unraveling the genetic etiology of Parkinson disease in a genomic era. Trends Genet *31*, 140-149.
Volkmann, J., Daniels, C., and Witt, K. (2010). Neuropsychiatric effects of subthalamic neurostimulation in Parkinson disease. Nat Rev Neurol *6*, 487-498.
von Campenhausen, S., Bornschein, B., Wick, R., Botzel, K., Sampaio, C., Poewe, W., Oertel, W., Siebert, U., Berger, K., and Dodel, R. (2005). Prevalence and incidence of Parkinson's disease in Europe. Eur Neuropsychopharmacol *15*, 473-490.
Wang, L., Park, H.J., Dasari, S., Wang, S., Kocher, J.P., and Li, W. (2013). CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model. Nucleic Acids Res *41*, e74.
Wang, S., Zhang, X., Guo, Y., Rong, H., and Liu, T. (2017). The long noncoding RNA HOTAIR promotes Parkinson's disease by upregulating LRRK2 expression. Oncotarget *8*, 24449-24456.
Wang, X., Song, X., Glass, C.K., and Rosenfeld, M.G. (2011). The long arm of long noncoding RNAs: roles as sensors regulating gene transcriptional programs. Cold Spring Harb Perspect Biol *3*, a003756.
Wooten, G.F., Currie, L.J., Bovbjerg, V.E., Lee, J.K., and Patrie, J. (2004). Are men at greater risk for Parkinson's disease than women? J Neurol Neurosurg Psychiatry *75*, 637-639.
Yan, W., Chen, Z.Y., Chen, J.Q., and Chen, H.M. (2018). LncRNA NEAT1 promotes autophagy in MPTP-induced Parkinson's disease through stabilizing PINK1 protein. Biochem Biophys Res Commun *496*, 1019-1024.
Zeighami, Y., Fereshtehnejad, S.M., Dadar, M., Collins, D.L., Postuma, R.B., Misic, B., and Dagher, A. (2017). A clinical-anatomical signature of Parkinson's disease identified with partial least squares and magnetic resonance imaging. Neuroimage.
Zhang, B., Gunawardane, L., Niazi, F., Jahanbani, F., Chen, X., and Valadkhan, S. (2014). A novel RNA motif mediates the strict nuclear localization of a long noncoding RNA. Mol Cell Biol *34*, 2318-2329.
Zuccato, C., Tartari, M., Crotti, A., Goffredo, D., Valenza, M., Conti, L., Cataudella, T., Leavitt, B.R., Hayden, M.R., Timmusk, T.*, et al.* (2003). Huntingtin interacts with REST/NRSF to modulate the transcription of NRSE-controlled neuronal genes. Nat Genet *35*, 76-83.
Zuo, L.J., Yu, S.Y., Hu, Y., Wang, F., Piao, Y.S., Lian, T.H., Yu, Q.J., Wang, R.D., Li, L.X., Guo, P.*, et al.* (2016). Serotonergic dysfunctions and abnormal iron metabolism: Relevant to mental fatigue of Parkinson disease. Sci Rep *6*, 19.

Identification et caractérisation de longs ARN non codants exprimés dans les neurones dopaminergiques et potentiellement impliqués dans la maladie de Parkinson

Résumé :

La maladie de Parkinson est une pathologie neurodégénérative caractérisée par des symptômes moteurs essentiellement causés par la dégénérescence des neurones dopaminergiques (DA) de la substance noire *pars compacta* (SN*pc*). Des analyses sur de larges cohortes de patients ont démontré le rôle déterminant des facteurs génétiques dans l'étiologie de cette pathologie. Cependant, seuls 10% des cas de la maladie de Parkinson (MP) correspondent à des formes familiales monogéniques, la majorité des cas sporadiques restent inexpliqués. Un nombre croissant d'études ont montré le rôle d'éléments non-codants du génome dans des pathologies humaines comme le diabète ou les cancers, ouvrant une nouvelle voie de recherche notamment pour les maladies neurodégénératives. De plus, les longs ARN non codants (lncARN) constituent des répertoires présentant une spécificité cellulaire très importante et pourraient donc rendre compte de mécanismes moléculaires encore inconnus impliqués dans les pathologies associées à des types cellulaires précis.

Dans ce contexte, nous nous sommes intéressés particulièrement aux lncARN exprimés dans les neurones dopaminergiques, correspondant au type cellulaire qui dégénère dans la MP, et à leur contribution dans cette maladie.

Les lncARN étant très peu conservés entre les espèces, nous avons généré des neurones DA humains à partir de deux modèles cellulaires, des cellules pluripotentes induites humaines (hiPSC) et des cellules immortalisées LUHMES provenant de cellules embryonnaires humaines du mésencéphale ventral. Après tri cellulaire des neurones DA issus d'hiPSC à l'aide d'un rapporteur fluorescent sous le contrôle du promoteur de la tyrosine hydroxylase, nous avons établi le premier répertoire de lncARN exprimés dans les neurones DA humains par des techniques de séquençage à haut débit (RNA-seq). Ainsi, nous avons identifié 882 lncARN dont 142 nouveaux transcrits non annotés. Ensuite, par comparaison avec des données génétiques obtenues par études d'association pangénomique, nous avons identifié 27 lncARN associés à des facteurs de risques de la MP. Nous avons montré l'expression dynamique de 9 d'entre eux au cours de la différenciation d'hiPSC en neurones DA. Finalement nous avons construit les outils moléculaires nécessaires à l'étude fonctionnelle de trois lncARN candidats, tous associés à au moins un facteur de risque pour la MP : *MIR4697HG, LNC122 et RP11-67M1.1*.

Ainsi, ce travail constitue une base importante pour évaluer la contribution et la fonction des lncARN dans la MP.

Mots clés :

Parkinson, longs ARN non-codants (lncARN), différenciation, cellules pluripotentes induites humaines (hiPSC), neurones dopaminergiques (DA), transcriptome.