



**HAL**  
open science

# Intérêt du transcriptome de cellules mononucléées périphériques sanguines dans l'étude des mécanismes moléculaires de la maladie de Parkinson

Aurore Nkiliza

► **To cite this version:**

Aurore Nkiliza. Intérêt du transcriptome de cellules mononucléées périphériques sanguines dans l'étude des mécanismes moléculaires de la maladie de Parkinson. Médecine humaine et pathologie. Université du Droit et de la Santé - Lille II, 2015. Français. NNT : 2015LIL2S058 . tel-02050189

**HAL Id: tel-02050189**

**<https://theses.hal.science/tel-02050189>**

Submitted on 27 Feb 2019

**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.

**UNIVERSITE DU DROIT ET DE LA SANTE DE LILLE II**  
**FACULTE DE MEDECINE**

**THESE DE DOCTORAT D'UNIVERSITE**  
**En Génétique, Génomique, Bioinformatique, Spécialité Neurosciences**

**Présentée par**  
**Aurore Nkiliza**

**Intérêt du transcriptome de cellules mononucléées périphériques  
sanguines dans l'étude des mécanismes moléculaires de la maladie  
de Parkinson**

Présentée le 11 décembre 2015, devant le jury composé de :

**Monsieur le Professeur Mathieu Anheim**

**Monsieur le Professeur Rejko Krüger**

**Madame le Docteur Marie-Christine Chartier-Harlin**

**Monsieur le Professeur Alain Destée**

**Rapporteur**

**Rapporteur**

**Présidente du jury**

**Directeur de thèse**

## **REMERCIEMENTS**

### ***A Monsieur le Professeur Destée,***

*Qui dès notre rencontre a parié sur mes capacités à relever les multiples obstacles qui jalonnent le parcours du doctorant, m'a accueillie et supportée matériellement dans son laboratoire et m'a permis de m'épanouir dans le domaine des neurosciences qui jusqu'ici m'était totalement étranger. En témoignage de ma profonde reconnaissance.*

### ***A Madame le Docteur Chartier-Harlin,***

*Qui m'a accueillie dans son groupe de recherche, a suivi avec intérêt le développement de cette thèse, m'a fait bénéficier de son expérience et de ses grandes compétences dans le domaine de la génétique des maladies neurodégénératives et m'a apportée une aide matériel inoubliable. En témoignage de ma vive gratitude et de mon profond respect.*

### ***A Monsieur le Professeur Anheim,***

*Qui a répondu présent, s'est montré disponible et a accepté de prendre part à la lecture de ce travail. Je vous adresse mes remerciements les plus sincères.*

### ***A Monsieur le Professeur Krüger,***

*Qui a manifesté sa volonté et son enthousiasme pour les travaux réalisés au cours de cette thèse et qui a accepté d'assister à leur présentation. Veuillez accepter mes plus profonds remerciements.*

### ***A Monsieur le Docteur Figeac, A Monsieur le Docteur Leprêtre, A Christophe Demay et A Céline Villenet***

*Pour les nombreuses heures passées à générer et traiter ces données transcriptomiques qui sans vous n'auraient jamais pu voir le jour. Vous avez su vous montrer disponibles pour répondre à mes questions et fait montre d'une pédagogie et d'une gentillesse admirable.*

*En témoignage de ma reconnaissance.*

### ***A Madame le Docteur Vanbesien-Maillot***

*Qui m'a écoutée et conseillée pendant toutes ces années et auprès de qui j'ai pu apprendre tant sur le plan scientifique que sur le plan humain. Malgré toutes vos responsabilités, vous avez toujours pris le temps de considérer et de répondre à mes interrogations. En témoignage de ma profonde admiration.*

### ***A Madame le Docteur Viltart***

*Pour vos encouragements et votre bonne humeur.*

**A Aliçia Stievenard**

*Sans qui ces trois années n'auraient tout simplement pas été les mêmes. Tu as su m'écouter, m'épauler et partager avec moi cette expérience marquante qu'est la thèse.*

*En témoignage de mon amitié sincère.*

**A Pierre Semaille et Thomas Comptdaer**

*Qui ont su me supporter pendant ces trois années, moi et ce Sud chère à mon cœur. Nous avons partagé ensemble bon nombre de moments qui m'auront marquée et grâce à vous je repartirai vers le Sud en me disant que finalement le Nord n'est peut-être pas si mal...*

*En témoignage de mon affection.*

**A tous les étudiants, post-doctorants et chercheurs de l'équipe**

*Avec qui j'ai partagé plus qu'une expérience professionnelle.*

**A Lucie Hopes**

*Sans qui je n'aurais jamais pu faire cette thèse. C'est grâce à ta gentillesse, ton ouverture d'esprit et ta pédagogie que les portes de l'accès à la thèse m'ont été ouvertes. En témoignage de ma gratitude.*

**A tous les membres du personnel de l'IRCL et du centre JParC**

*Pour leur gentillesse, leur joie de vivre et toutes ces discussions qui hanteront à jamais les couloirs de l'IRCL.*

**A mon amie Cacylde,**

*Que je ne remercierai jamais assez pour avoir été présente pendant ces trois années.*

**A toi David,**

*Qui malgré la distance a su me montrer que tu étais toujours présent et là pour moi.*

**Et enfin, à ma famille,**

*Vous avez toujours fait montre d'un soutien sans faille. Sans votre écoute et vos encouragements, je n'aurais jamais pu mener cette thèse à bien. Vous avez vécu avec moi toutes les étapes qui m'auront menée jusqu'ici et c'est grâce à vous que j'en suis là aujourd'hui. Les mots ne sauraient dire à quel point je vous suis reconnaissante et fière de vous avoir à mes côtés.*

*Hamwe nanjye ibihe byose.*

*« Ntakigerwaho kidakorewe »*

Ce travail a donné lieu aux publications et communications suivantes :

### Publications

Mutez E., **Nkiliza A.**, Belarbi K., de Broucker A., Vanbesien-Mailliot C., Bleuse S., Duflot A., Comptdaer T., Semaille P., Blervaque R., Hot D., Leprêtre F., Figeac M., Destée A., Chartier-Harlin M.C. (2014). Involvement of the immune system, endocytosis and EIF2 signaling in both genetically determined and sporadic forms of Parkinson's disease. *Neurobiology of Disease*, 63, 165–170.

**Nkiliza A.**, Mutez E., Simonin C., Leprêtre F., Duflot A., Figeac M., Genet A., Sablonnière B., Devos D., Defebvre L., Destée A., Chartier-Harlin M.C. (2015). RNA binding disturbances as a continuum from Spinocerebellar ataxia type 2 to Parkinson's disease (en soumission à *Brain*).

Taymans J-M., **Nkiliza A.**, Chartier-Harlin M.C. (2015). Deregulation of protein translation control, a potential game-changing hypothesis for Parkinson's disease pathogenesis. *Trends in Molecular Medicine*, 21(8), 466–472.

### Communications affichées

**A. Nkiliza**, L. Hopes, C. Demay, E. Mutez, F. Leprêtre, S. Bleuse, A. Destée, M. Figeac, M.C. Chartier-Harlin «Study of RNA splicing by RNAseq in Parkinson's disease». Journée des doctorants, Lille 2013.

**A. Nkiliza**, L. Hopes, C. Demay, E. Mutez, F. Leprêtre, S. Bleuse, A. Destée, M. Figeac, M.C. Chartier-Harlin «Study of RNA splicing by RNAseq in Parkinson's disease». 11<sup>ème</sup> Colloque de la Société des neurosciences, Lyon 2013.

**A. Nkiliza**, L. Hopes, C. Demay, E. Mutez, F. Leprêtre, S. Bleuse, A. Destée, M. Figeac, M.C. Chartier-Harlin «Study of RNA splicing by RNAseq in Parkinson's disease». 14<sup>ème</sup> édition des Journées Ouvertes en Biologie, Informatique & Mathématiques, Toulouse 2013.

M.C. Chartier-Harlin, **A. Nkiliza**, L. Hopes, E. Mutez, S. Bleuse, C. Demay, M. Figeac, A. Destée «Perturbation of RNA splicing variants in Parkinson's disease». 12th International Meeting of the Cellular aspects of mRNA fate, Paris 2013.

**A. Nkiliza**, S. Maniez, L. Hopes, C. Demay, E. Mutez, T. Comptdaer, P. Semaille, S. Bleuse, F. Leprêtre, M. Figeac, R. Polakowska, A. Destée, M.C. Chartier-Harlin «Identification of RNA splicing perturbations in Parkinson's disease». 14<sup>ème</sup> édition de la Journée André Verbert, Lille 2014.

**A. Nkiliza**, E. Mutez, C. Simonin, A. Duflot, M. Figeac, B. Sablonnière, D. Devos, L. Defebvre, A. Destée, M.C. Chartier-Harlin «RNA metabolism disturbances in genetic and sporadic forms of Parkinson's disease » Colloque Dementia in Neurological and Mental diseases, Lille 2015.

## Table des matières

1.	Introduction.....	5
1.1.	La maladie de Parkinson.....	5
1.1.1.	L’historique de la maladie de Parkinson .....	5
1.1.2.	La maladie de Parkinson : une maladie à prédominance motrice .....	7
1.1.2.1.	Les symptômes moteurs .....	7
1.1.2.1.1.	Le tremblement de repos.....	7
1.1.2.1.2.	La bradykinésie.....	7
1.1.2.1.3.	Le freezing .....	7
1.1.2.1.4.	La rigidité.....	8
1.1.2.1.5.	L’instabilité posturale.....	8
1.1.2.2.	L’anatomopathologie de la maladie de Parkinson .....	8
1.1.2.2.1.	Le rôle des ganglions de la base.....	8
1.1.2.2.2.	La dégénérescence neuronale.....	10
1.1.2.2.3.	Les agrégats protéiques : corps et neurites de Lewy .....	11
1.1.2.3.	Le diagnostic clinique .....	11
1.1.2.4.	Les traitements.....	12
1.1.2.4.1.	Les traitements médicamenteux.....	12
1.1.2.4.2.	Les traitements chirurgicaux .....	14
1.1.2.4.3.	La vaccination .....	15
1.1.3.	La maladie de Parkinson : une maladie plus complexe.....	15
1.1.3.1.	Les symptômes non moteurs .....	15
1.1.3.1.1.	Les troubles neuro-psychiatriques .....	16
1.1.3.1.2.	Les troubles du sommeil.....	17
1.1.3.1.3.	Les troubles sensoriels .....	18
1.1.3.1.4.	Les troubles du système nerveux autonome .....	18
1.1.3.1.5.	Les troubles gastro-intestinaux .....	18
1.1.3.2.	Les signes anatomopathologiques : la réponse inflammatoire.....	19
1.1.4.	Les mécanismes moléculaires contribuant au développement de la MP.....	20
1.1.4.1.	L’apport des facteurs environnementaux .....	20
1.1.4.2.	L’apport des facteurs génétiques.....	24
1.1.4.2.1.	Le gène <i>SNCA</i> .....	26
1.1.4.2.2.	Le gène <i>LRRK2</i> .....	30
1.1.4.2.3.	Le gène <i>ATXN2</i> .....	33
1.1.4.2.4.	Le gène <i>PRKN</i> .....	37
1.1.4.2.5.	Le gène <i>PINK1</i> .....	38



1.1.4.2.6. Le gène <i>DJ-1</i> .....	39
1.1.4.2.7. Le gène <i>VPS35</i> .....	40
1.1.4.2.8. Le gène <i>EIF4G1</i> .....	41
1.1.4.3. L'apport des études d'association.....	43
1.1.4.4. La pertinence de l'étude du transcriptome dans la maladie de Parkinson.....	45
1.1.4.5. L'apport du transcriptome de cerveaux <i>post-mortem</i> .....	47
1.1.4.6. L'apport du transcriptome de cellules dérivées de patients.....	49
1.1.5. Les perturbations du métabolisme des ARNs dans les maladies neurodégénératives : le rôle de l'épissage.....	51
1.1.5.1. La définition de l'épissage .....	51
1.1.5.2. La machinerie effectrice de l'épissage : le spliceosome.....	53
1.1.5.3. La réaction d'épissage .....	56
1.1.5.4. La régulation de l'épissage .....	57
1.1.5.5. Les maladies neurodégénératives : des exemples de perturbations de l'épissage.....	60
1.1.5.6. Les méthodes d'étude de l'épissage .....	61
1.1.5.6.1. Les puces d'épissage.....	62
1.1.5.6.2. Le séquençage des ARNs .....	63
2. Objectifs de la thèse .....	68
3. Résultats .....	71
3.1. <b>Chapitre 1</b> : Étude du transcriptome de cellules mononucléées périphériques sanguines de sujets porteurs de la mutation G2019S de <i>LRK2</i> et de sujets sporadiques. ....	71
<b>Article 1: Involvement of the immune system, endocytosis and EIF2 signaling in both genetically determined and sporadic forms of Parkinson's disease</b> .....	74
3.2. <b>Chapitre 2</b> : Étude du transcriptome de cellules mononucléées périphériques sanguines de sujets porteurs de mutations du gène <i>ATXN2</i> .....	98
<b>Article 2: RNA binding disturbances as a continuum from Spinocerebellar ataxia type 2 to Parkinson's disease</b> .....	101
3.3. <b>Chapitre 3</b> : Mise en évidence de perturbations de l'épissage dans la maladie de Parkinson.....	102
3.3.1. Méthodes .....	135
3.3.1.1. Sujets .....	135
3.3.1.2. Extraction d'ARNs à partir de cellules mononucléées périphériques sanguines .....	137
3.3.1.2.1. Protocole d'extraction des ARNs destinés au séquençage par le séquenceur SOLiD™ 4 « Sequencing by Oligonucleotide Ligation and Detection » .....	137
3.3.1.2.2. Protocole d'extraction des ARNs destinés au séquençage par le séquenceur Ion Proton™ .....	138
3.3.1.3. Préparation de la librairie d'ADNc à séquencer .....	138
3.3.1.3.1. Préparation de la librairie pour le séquençage par le séquenceur SOLiD™ 4 .....	139

3.3.1.3.2. Préparation de la librairie pour le séquençage par le séquenceur Ion Proton™ .....	140
3.3.1.4. Réaction de séquençage.....	140
3.3.1.4.1. Séquençage effectué par le séquenceur SOLiD™ 4.....	140
3.3.1.4.2. Séquençage effectué par le séquenceur Ion Proton™ .....	142
3.3.1.5. Analyse quantitative des défauts d'épissage .....	142
3.3.1.5.1. A partir des données de séquençage SOLiD™ 4 .....	142
3.3.1.5.2. A partir des données de séquençage Ion Proton™ .....	143
3.3.1.6. Analyse qualitative des défauts d'épissage.....	143
3.3.1.6.1. A partir des données de séquençage SOLiD™ 4 .....	143
3.3.1.6.2. A partir des données de séquençage Ion Proton™ .....	145
3.3.1.6.3. A partir des données combinées de séquençage SOLiD™ 4 et Ion Proton™ .....	146
3.3.1.7. Caractérisation des processus biologiques, des voies canoniques et des gènes dérégulés .....	146
3.3.1.7.1. Logiciel Gene Set Enrichment Analysis (GSEA).....	146
3.3.1.7.2. Logiciel Ingenuity Pathways Analysis (IPA®).....	147
3.3.1.7.3. Création de liste de gènes .....	147
3.3.1.8. Amplification par PCR, RT-PCR et RT-qPCR .....	148
3.3.2. Résultats .....	151
3.3.2.1. Étude de faisabilité (séquençage SOLiD™ 4) .....	151
3.3.2.1.1. Analyse quantitative de l'expression des gènes.....	151
3.3.2.1.2. Analyse quantitative de l'expression des transcrits.....	154
3.3.2.1.3. Analyse qualitative de l'épissage : création de nouvelles jonctions entre exons .....	156
3.3.2.2. Étude pilote (séquençage Ion Proton™).....	161
3.3.2.2.1. Analyse quantitative de l'expression des gènes.....	161
3.3.2.2.2. Analyse quantitative de l'expression des transcrits.....	163
3.3.2.2.3. Analyse qualitative de l'épissage : création de nouvelles jonctions entre exons .....	166
3.3.2.3. Étude des données combinées de séquençage SOLiD™ 4 et Ion Proton™ .....	167
3.3.2.3.1. Analyse qualitative de l'épissage : création de nouvelles jonctions entre exons .....	167
3.3.3. Discussion .....	171
3.4. <b>Chapitre 4 : Rôle potentiel de la traduction dans la maladie de Parkinson.....</b>	175
<b>Article 3: Deregulation of protein translation control, a potential game-changing hypothesis for Parkinson's disease pathogenesis .....</b>	176
4. Discussion générale .....	199
4.1. Études des mécanismes moléculaires de la maladie de Parkinson.....	199
4.2. Importance de la voie de signalisation EIF2 dans la maladie de Parkinson : son lien avec la traduction et l'épissage .....	201

4.3.	Perturbations du métabolisme des ARNs : Implication des mécanismes de dégradation .....	206
4.4.	L'épissage alternatif comme source de biomarqueurs potentiels de la maladie de Parkinson.....	209
5.	Conclusion .....	212
6.	Annexes .....	215
	<b>Annexe 1: Critères diagnostiques UKPDSBB de la maladie de Parkinson (United Kingdom Parkinson's Disease Society Brain Bank) .....</b>	<b>215</b>
	<b>Annexe 2: Échelle de Hoehn et Yahr .....</b>	<b>217</b>
	<b>Annexe 3 : Échelle d'évaluation unifiée pour la maladie de Parkinson (UPDRS, Unified Parkinson Disease Rating Scale).....</b>	<b>218</b>
	<b>Annexe 4 : Critères de diagnostic clinique du NINDS (National Institute of Neurological Disorders and Stroke) .....</b>	<b>231</b>
7.	Références.....	234

## INDEX DES FIGURES

**Figure 1: Organisation des noyaux gris centraux**

**Figure 2: Circuit moteur des ganglions de la base**

**Figure 3: Coupe transversale de cerveaux de sujets parkinsoniens et de sujets sains**

**Figure 4: Agrégats protéiques caractéristiques de la maladie de Parkinson**

**Figure 5: Métabolisme de la dopamine et sites d'action des traitements médicamenteux**

**Figure 6: Progression de la pathologie  $\alpha$ -synucléine selon l'hypothèse de Braak**

**Figure 7: Synthèse et acheminement du MPP<sup>+</sup> vers les neurones dopaminergiques de la *SNpc***

**Figure 8: Similarité de structure chimique entre le MPTP, le MPP<sup>+</sup> et le paraquat**

**Figure 9: Biosynthèse de l'ATP à travers la chaîne respiratoire mitochondriale**

**Figure 10: Localisation des mutations et des domaines protéiques de l' $\alpha$ -synucléine**

**Figure 11: Représentation de la protéine LRRK2 et des mutations causales de la maladie de Parkinson**

**Figure 12: Domaines protéiques et répétitions de polyglutamine dans la protéine ataxine-2**

**Figure 13: Domaines protéiques de la parkin**

**Figure 14: Représentation de la protéine EIF4G1 et de ses mutations**

**Figure 15: Mécanismes moléculaires physiopathologiques de la maladie de Parkinson**

**Figure 16: Différents types d'ARNs codés par le génome humain**

**Figure 17: Principe de l'analyse du transcriptome à l'aide de puce d'expression**

**Figure 18: Les différentes étapes de l'expression génique**

**Figure 19: Diversité des transcrits générés par épissage alternatif**

**Figure 20: Composition des 5 sous-unités du spliceosome**

**Figure 21: Assemblage des snRNPs**

**Figure 22: Motifs des sites d'épissage**

**Figure 23: Principe de la réaction d'épissage**

**Figure 24: Réaction d'épissage**

**Figure 25: Régulation de l'épissage par phosphorylation des protéines SR**

**Figure 26: Analyse des événements d'épissage par l'utilisation de micropuces**

**Figure 27: Principe du RNAseq**

**Figure 28: Préparation de la librairie pour le séquençage par le séquenceur SOLiD™ 4**

**Figure 29: Principe du séquençage par le séquenceur SOLiD™ 4**

**Figure 30: Critères de détection des jonctions entre exons par le logiciel Bioscope®**

**Figure 31: Hybridation d'amorces oligonucléotides de PCR sur la jonction nouvelle**

**Figure 32: Moyennes des ratios EA/T et P/T chez les témoins et chez les malades**

**Figure 33: Amplicons de RT-PCR après dépôt sur gel d'agarose**

**Figure 34: Quantification par RT-qPCR des transcrits de *RBM5* et *CCNL1***

## INDEX DES TABLEAUX

Tableau 1: Loci PARK liés à la maladie de Parkinson

Tableau 2: Synthèse des avantages du RNAseq par rapport aux puces d'épissage

Tableau 3: Sélection de voies canoniques associées aux gènes différentiellement exprimés dans nos données et dans la littérature

Tableau 4: Sélection de gènes de la voie EIF2 dérégulés chez les parkinsoniens porteurs de la mutation G2019S de *LRRK2* et chez les sujets sporadiques

Tableau 5: Processus biologiques communs aux parkinsoniens sporadiques, porteurs de la mutation G2019S de *LRRK2* ou d'expansions de triplets CAG dans *ATXN2*

Tableau 6: Sujets recrutés pour l'étude de l'épissage par séquençage des ARNs

Tableau 7: Différents types de transcrits identifiés après reconstitution par Cufflinks.

Tableau 8: Amorces oligonucléotides utilisées pour les PCR, RT-PCR et RT-qPCR

Tableau 9: Nombre de reads générés par individu et alignés sur le génome humain RefSeq hg19

Tableau 10: Top 15 des processus biologiques liés aux gènes dérégulés

Tableau 11: Gènes de facteurs liés à l'épissage différentiellement exprimés entre les parkinsoniens et les témoins

Tableau 12: Top 15 des processus biologiques liés aux transcrits dérégulés

Tableau 13: Transcrits de gènes liés à l'épissage différentiellement exprimés entre les malades et les témoins

Tableau 14: Jonctions nouvelles validées par RT-PCR

Tableau 15: Voies canoniques communes aux différentes analyses du transcriptome

Tableau 16: Nombre de reads par sujet, alignés sur le génome humain Ensembl GRCh37

Tableau 17: Processus biologiques liés aux gènes dérégulés chez les parkinsoniens à stades précoces

Tableau 18: Douzes facteurs liés à l'épissage dérégulés chez les parkinsoniens

Tableau 19: Top 18 des processus biologiques liés aux transcrits dérégulés

Tableau 20: Transcrits de gènes liés à l'épissage

**Tableau 21: Vingt-huit transcrits à jonctions nouvelles, dérégulés entre les malades et les témoins**

**Tableau 22: Top 20 des processus biologiques liés aux gènes à défauts d'épissage**

**Tableau 23: Gènes de facteurs liés à l'épissage ayant des perturbations de leur épissage**

**Tableau 24: Gènes de la maladie de Parkinson avec un Sd significatif**

## LISTE DES ABBREVIATIONS

3'UTR	3' Untranslated Région
4E-BP	Eukaryotic translation initiation factor 4E binding protein
6-OHDA	6-Hydroxy-Dopamine
ADNg	Acide DésoxyriboNucléique génomique
ALS	Amyotrophic lateral sclerosis
ATP	Adénosine Triphosphate
ATXN2	Ataxine 2
COX2	Cyclooxygenase 2
DCL	Démence à corps de Lewy
DNMT1	DNA (Cytosine-5-)-Methyltransferase 1
GABA	Acide $\gamma$ -aminobutyrique
GATA1	GATA binding protein 1
GATA2	GATA binding protein 2
GPe	Globus pallidus externe
GPi	Globus pallidus interne
GSEA	Gene Set Enrichment Analysis
H2O2	Peroxyde d'hydrogène
hnRNP	Heterogeneous ribonucleoprotein particle
IPA	Ingenuity® Pathway Analysis
L-Dopa	Levodopa
LRRK2	Leucine-rich repeat kinase 2
MAOA	Monoamine oxidase A
MAOB	Monoamine oxidase B
miARNs	micro-ARNs
MP	Maladie de Parkinson



MPP+	1-Methyl-4-phenylpyridinium ion
MPTP	1-méthyl-4-phényl-1,2,3,6-tétrahydropyridine
NMD	Nonsense-mediated mRNA decay
OMIM	Online Mendelian Inheritance in Man
PABP	Poly(A)-binding protein
PABPC1	PolyA-Binding Protein Cytoplasmic 1
PAM2	PABP-interacting motif
PD	Parkinson's disease
PGC1- $\alpha$	PPARGC1A
PINK1	PTEN-induced putative kinase 1
polyA	Polyadénylation
polyQ	Polyglutamine
protéines SR	Protéines riches en sérines et arginines
RNAseq	RNA sequencing
RNS	Reactive nitrogen species
ROS	Reactive oxygen species
SCA2	Spinocerebellar Ataxia Type 2
SCA2c	Patient porteur de mutation de l' <i>ATXN2</i> avec un phénotype cérébelleux
SCA2p	Patient porteur de mutation de l' <i>ATXN2</i> avec un phénotype parkinsonien
SLA	Sclérose latérale amyotrophique
SNAP-25	Synaptosomal-associated protein 25
SNCA	Synuclein, alpha
SNpc	Substantia nigra pars compacta
SNr	Substantia nigra reticulata
snRNPs	small nuclear Ribonucleic proteins
SOD2	Superoxide dismutase 2
VAMP	Vesicle-associated membrane protein

VPS35

Vacuolar Protein Sorting 35

## Préambule

Au cours du XXème siècle, l'espérance de vie a considérablement augmentée et s'est accompagnée d'une hausse des maladies liées au vieillissement. L'âge est un facteur de risque majeur des maladies neurodégénératives dont la seconde plus fréquence est la maladie de Parkinson, qui fait l'objet de cette thèse. D'abord considérée comme une maladie du système nerveux centrale, elle a principalement été associée à une perte de neurones dopaminergiques et à la formation d'agrégats protéiques nommés corps de Lewy dans la substance noire. Ces lésions sont à l'origine d'un déficit en dopamine qui se manifeste par l'apparition de divers troubles moteurs tels que le tremblement, la rigidité, le ralentissement des mouvements ou encore la perte d'équilibre. Ces manifestations cliniques, fortement handicapantes, sont prises en charge par diverses thérapeutiques qui pallient au déficit en dopamine et dont les effets bénéfiques s'estompent au cours du temps. Si les troubles moteurs constituent les signes les plus visibles, l'identification de lésions au niveau du système nerveux autonome et des troubles non moteurs peu sensibles à la dopa-thérapie, a changé la façon de concevoir la maladie. Ces observations ont permis de redessiner son tableau clinique et de constater que sa physiopathologie était plus complexe. Cependant, malgré ces avancées cliniques, la maladie de Parkinson demeure toujours une maladie incurable. Le manque de traitement curatif réside en partie dans la méconnaissance des mécanismes moléculaires qui sous-tendent la physiopathologie de la maladie, limitant l'identification de marqueurs cliniques et de cibles thérapeutiques. Une amélioration des connaissances de ces mécanismes apparaît alors indispensable à une meilleure prise en charge.

La mise en évidence de facteurs environnementaux et génétiques à l'origine de certaines formes de la maladie de Parkinson a marqué un second tournant dans la recherche des déterminants à l'origine de la maladie. Les études de ces facteurs environnementaux et de ces formes génétiques rares ont permis d'identifier plusieurs mécanismes moléculaires impliqués dans la physiopathologie. Ces déterminants ont été d'autant plus importants qu'ils ont révélé des changements moléculaires communs aux sujets présentant des formes sporadiques majoritaires. Ainsi, l'étude des formes génétiques apparaît être un bon moyen d'identifier les mécanismes moléculaires physiopathologiques d'un grand nombre de cas de la maladie.

Les approches de génétique classique et de génétique inverse qui ont grandement contribué à l'identification fonctionnelle des gènes mutés, présentent cependant certaines limites. Elles ne permettent qu'une identification fonctionnelle ciblée limitant ainsi l'étendue des processus cellulaires étudiés et la connaissance de perturbations moléculaires associées à la maladie. Une approche plus globale semble être le moyen le plus efficace d'identifier l'ensemble des mécanismes impactés par cette maladie complexe.

La complexité structurale et fonctionnelle du cerveau est régentée par de nombreux gènes qui assurent le bon fonctionnement et l'homéostasie cérébrale. L'expression génique se découpe en plusieurs étapes allant de la transcription à la synthèse protéique, en passant par les diverses étapes de maturation des ARNs. Des modifications du profil d'expression génique ont été suggérées comme étant à l'origine de la dégénérescence neuronale. Ainsi, l'observation des profils d'expression génique des patients atteints de la maladie de Parkinson devrait permettre de mieux comprendre les modifications à l'origine de la mort neuronale. Traditionnellement, l'étude des profils d'expression génique passe par une analyse du transcriptome correspondant à l'ensemble des transcrits codés par le génome (ARNs). Ces études permettent d'une part l'étude des niveaux d'expression des différents transcrits d'un gène et d'autre part une étude de la structure de ces différents transcrits. Pendant de nombreuses années, les micropuces ont été un outil de choix dans l'étude des changements d'expression de gènes. Par l'analyse simultanée de plusieurs échantillons, elles ont permis d'effectuer des analyses d'expression différentielle mettant en évidence des modifications de l'expression d'un grand nombre de gènes dont certaines sont spécifiques et caractéristiques d'un échantillon donné.

La découverte de la complexité du transcriptome a conduit aux développements de technologies permettant d'aller au-delà de l'analyse des niveaux d'expression des ARNs. En effet, les transcrits nouvellement synthétisés sont interrompus par des morceaux de séquences non codantes appelées introns, les séquences codantes étant appelées exons. Au cours de la maturation des ARNs (transcrits), ces introns sont excisés et les exons sont rattachés entre eux : c'est le processus d'épissage. Pour un même gène, plusieurs événements d'épissages différents peuvent être réalisés et conduire à l'exclusion d'un ou plusieurs exons : on parle alors d'épissage alternatif. Ces épissages alternatifs conduisent, pour un même gène, à la production de plusieurs ARN messagers qui donneront des protéines aux propriétés sensiblement différentes. Cet épissage alternatif apporte ainsi une possibilité de diversification et de contrôle des propriétés d'une protéine donnée très large. Les avancées technologiques en matière de séquençage et de traitement des données à haut débit ont conduit à améliorer les méthodes d'exploration des profils d'expression génique, en partie par le développement des techniques de micropuces d'épissage et de séquençage des ARNs. Basées sur l'analyse de l'ensemble des transcrits, ces techniques offrent la possibilité d'une large exploration des variations au niveau des transcrits contenus dans un échantillon donné. Ainsi certaines puces permettent une analyse de l'expression tandis que d'autres permettent d'analyser la structure de ces transcrits (compositions en exons) et donc d'analyser des variations qualitatives tels que l'épissage. Bien qu'informatives ces puces sont basées sur un principe d'appariement des transcrits à des sondes complémentaires, ce qui limite l'étude aux transcrits ayant une sonde correspondante. A l'inverse, le

séquençage des ARNs ne dépend pas de cet appariement et permet d'avoir accès à la séquence de l'ensemble des transcrits contenus dans un échantillon. Le développement de cette technique a donc élargi le champ des transcrits étudiés et permet d'observer des variations occultées par les micropuces.

Ces méthodes d'étude du transcriptome ont déjà fait leurs preuves dans leur capacité à définir une signature moléculaire caractéristique d'une pathologie donnée. Par l'analyse simultanée de variations quantitatives et qualitatives de l'ensemble des gènes et transcrits, elles offrent ainsi l'accès à un large panel d'informations élargissant les connaissances des variations spécifiques d'une condition donnée. Aux vues des capacités à identifier de nombreuses variations et afin d'avoir une vue d'ensemble des mécanismes moléculaires de la maladie de Parkinson, nous avons entrepris une étude du transcriptome de cellules mononucléées périphériques sanguines de sujets présentant des formes monogéniques et sporadiques de la maladie dans le but de mettre en évidence les perturbations cellulaires pouvant en être à l'origine. Ce travail qui a constitué l'essentiel de cette thèse est une preuve de concept de la pertinence de l'utilisation des cellules mononucléées périphériques sanguines comme modèle d'étude de la maladie permettant 1) de mettre en évidence des altérations de mécanismes déjà décrits comme étant impliqués dans la maladie, 2) de mettre l'accent sur de nouveaux mécanismes potentiels de la physiopathologie de la maladie, 3) d'étudier l'altération de ces mécanismes au cours de l'évolution afin de définir des altérations dès les premiers stades de la maladie qui pourraient être à l'origine de la mort neuronale et ciblées par de nouveaux traitements.

# **INTRODUCTION**

## 1. Introduction

### 1.1. La maladie de Parkinson

#### 1.1.1. L'histoire de la maladie de Parkinson

La maladie de Parkinson (MP) est la seconde maladie neurodégénérative la plus commune après la maladie d'Alzheimer. Les études épidémiologiques ont déterminé plusieurs facteurs de risques influant sur sa fréquence<sup>1</sup> : le sexe masculin<sup>2</sup>, l'ethnie (fréquence plus élevée chez les caucasiens)<sup>2,3,4</sup> et l'âge<sup>5</sup> ; ce dernier étant le facteur de risque prédominant. La prévalence de la maladie de 1% chez les sujets âgés de plus de 65 ans peut alors atteindre 5% chez les personnes de plus de 85 ans<sup>6</sup>. D'après une étude parue en 2007, le nombre de cas de MP rapportés en 2005 sera doublé d'ici 2030 du fait du vieillissement croissant de la population<sup>7</sup>. La MP constitue donc un problème de santé publique majeur.

Initialement décrite sous le nom de paralysie agitante par le médecin anglais James Parkinson, elle porte depuis les travaux de Jean-Martin Charcot le nom de ce médecin anglais. Dans son essai paru en 1817 « *An essay on the shaking palsy*<sup>8</sup> », il faisait la toute première description détaillée des symptômes présentés par les patients parkinsoniens : « *Ces sujets sont caractérisés par une faiblesse et un tremblement dans différentes parties du corps telles que la tête, les mains et les bras. Ces symptômes semblent s'aggraver avec l'évolution de la maladie et ce de manière plus importante dans la région où se sont déclarés les premiers symptômes. L'exécution des mouvements est de plus en plus difficile, l'écriture et la lecture étant affectées du fait du tremblement. Aussi, les difficultés d'exécution des mouvements du quotidien affectent de plus en plus la qualité de vie. Outre les symptômes moteurs, l'humeur des patients semble également être affectée, accompagnée d'un état de tristesse évoluant jours après jours. Des troubles du sommeil peuvent également être observés certains étant associés à des épisodes de tremblements maintenant le patient en éveil. Les troubles et le manque de sommeil peuvent engendrer des épisodes de délire. Certains patients présentent des problèmes de salivation excessive, de miction et vont à la selle de façon involontaire. Ainsi, l'évolution de la maladie est de plus en plus invalidante et dégradante pour les personnes qui en sont atteintes.* »

Si cette description mentionnait déjà des symptômes moteurs et non moteurs, pendant des dizaines d'années, seuls les symptômes moteurs ont bénéficié d'une attention particulière du fait de leur caractère prédominant. Le tremblement, la rigidité, la bradykinésie et l'instabilité posturale définissent ainsi une tétrade de symptômes qui à eux seuls ont longtemps défini le tableau clinique de la maladie. Cette tétrade corrèle avec la dégénérescence des neurones dopaminergiques de la

substance noire et la présence d'inclusions intra-cytoplasmiques nommées corps de Lewy dans les neurones persistants. Ces lésions sont à l'origine d'une baisse du niveau de dopamine striatale, qui nous le verrons, est importante dans le contrôle de la motricité. L'apport de dopamine par les dopa-thérapies est alors apparu comme le mode de traitement le plus adapté afin de répondre à ce déficit. Cependant, les bénéfices conférés par ces thérapies sont limités dans le temps (lune de miel), contrebalancés par l'apparition d'effets secondaires tels que les mouvements involontaires (dyskinésies) et les hallucinations<sup>9</sup>. De plus, leurs effets quasi-inexistants sur les symptômes non moteurs ont suggéré que l'origine lésionnelle de la maladie ne se limitait pas uniquement à la *SNpc*.

Les études cliniques de ces dix dernières années ont permis de reconsidérer l'importance des signes non moteurs dans la physiopathologie de la MP<sup>10,11,12</sup>. Ces symptômes qui jusque-là avaient été minimisés, corrèlent avec la présence de lésions anatomiques dans le système nerveux central et périphérique. Ces découvertes ont permis d'améliorer les connaissances de la maladie et de prendre conscience de sa complexité. A l'heure actuelle, ces symptômes non moteurs prennent une importance toute particulière puisqu'il est maintenant admis qu'ils apparaissent avant les troubles moteurs, laissant entrevoir la possibilité d'un diagnostic précoce<sup>13</sup>. Les traitements actuels agissant essentiellement sur la symptomatologie motrice, il est donc nécessaire de développer de nouveaux traitements qui permettraient une prise en charge de l'ensemble des symptômes et qui pourraient agir dès les premiers stades de la maladie. Toutefois, le développement de nouveaux traitements symptomatiques et/ou curatifs ayant des effets sur la progression de la maladie, requiert l'identification de nouvelles cibles thérapeutiques et de marqueurs diagnostiques précoces permettant d'administrer ces nouvelles thérapies au moment où elles seront le plus efficaces.

Le développement de ces marqueurs et nouveaux traitements reposent sur la connaissance des mécanismes moléculaires d'une maladie. Elle est nécessaire à l'identification de cibles potentielles et à la caractérisation de variations spécifiques pouvant servir d'aide au diagnostic. Depuis la découverte de la MP, la question des altérations moléculaires contribuant à son développement a été soulevée. Toutefois, c'est la découverte de déterminants environnementaux et génétiques qui a permis de proposer des mécanismes potentiels à l'origine de la mort neuronale qui, à l'heure actuelle, ne sont pas totalement compris.

C'est dans ce contexte que s'inscrit ce travail de thèse visant à apporter un éclairage sur les mécanismes à l'origine de la MP qui permettrait d'identifier des biomarqueurs précoces et aiderait au développement de nouvelles thérapies.



## 1.1.2. La maladie de Parkinson : une maladie à prédominance motrice

### 1.1.2.1. Les symptômes moteurs

Traditionnellement, les premiers signes visibles de la MP sont des symptômes moteurs caractéristiques qui définissent une tétrade parkinsonienne incluant le tremblement de repos, la bradykinésie, la rigidité et l'instabilité posturale sur laquelle repose le diagnostic.

#### 1.1.2.1.1. Le tremblement de repos

Le tremblement de repos est le symptôme moteur le plus communément observé, révélateur chez les trois quart des malades et présent chez 80% d'entre eux. Pour certains, il est même constamment présent à un moment ou à un autre de l'évolution de la maladie. Le tremblement est initialement unilatéral pour ensuite se bilatéraliser. Qualifié de tremblement de repos, il se manifeste lorsque le patient est inactif, apparaissant ou s'aggravant avec l'état émotionnel du patient. C'est un tremblement lent avec une fréquence de 4 à 6 Hertz (Hz). Le plus souvent présent à la partie distale des membres, principalement supérieurs, il est décrit comme un mouvement d'émission des doigts, de flexion/extension du poignet, de supination/pronation de l'avant-bras. Cependant certains patients se plaignent également de tremblement au niveau des membres inférieurs, des lèvres et de la mâchoire.

#### 1.1.2.1.2. La bradykinésie

Ce symptôme est constant, indispensable au diagnostic clinique de la MP. La bradykinésie est caractérisée par le retard à l'initiation (augmentation du temps de réaction), la lenteur d'exécution et l'hypométrie (diminution de l'amplitude) des mouvements et des gestes. Elle concerne les gestes volontaires en particulier alternatifs rapides (rasage, brossage des dents...) ou fins (boutonnage). Elle n'épargne pas les mouvements automatiques tels que le ballant du bras à la marche ou les mimiques (hypomimie donnant ainsi au patient une expression figée) et la déglutition. La marche est également concernée : lente, à petits pas avec plus tardivement des piétinements voire des blocages (freezing de la marche) au demi-tour ou au franchissement d'obstacle (passage d'une porte).

#### 1.1.2.1.3. Le freezing

Le freezing est une forme d'akinésie qui semble être plus fréquente chez l'homme que chez la femme et moins fréquente chez les sujets dont le premier symptôme est le tremblement. Il se définit comme le blocage d'une fonction (marche, ballant du bras, battement des paupières).

#### 1.1.2.1.4. La rigidité

La rigidité est une augmentation du tonus musculaire qui dans le cas de la MP (et des syndromes parkinsoniens) a pour caractéristique de prédominer sur les muscles fléchisseurs (expliquant la posture voûtée..) et d'être « plastique », le segment de membre conservant l'attitude que l'examineur lui a passivement imposé. Elle participe à la gêne motrice et aux douleurs chroniques.

#### 1.1.2.1.5. L'instabilité posturale

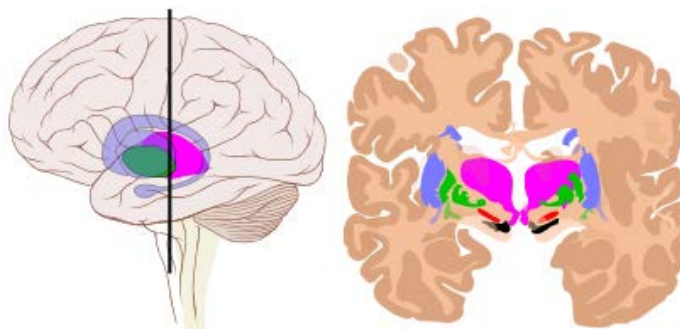
L'instabilité posturale apparaît plus tardivement et conduit à des chutes fréquentes et récurrentes qui mettent en jeu la sécurité du patient et peuvent le contraindre à limiter ses déplacements.

L'ensemble de ces symptômes moteurs résultent de la présence de lésions dans la *SNpc*, une région cérébrale située au centre du circuit moteur des ganglions de la base.

### 1.1.2.2. L'anatomopathologie de la maladie de Parkinson

#### 1.1.2.2.1. Le rôle des ganglions de la base

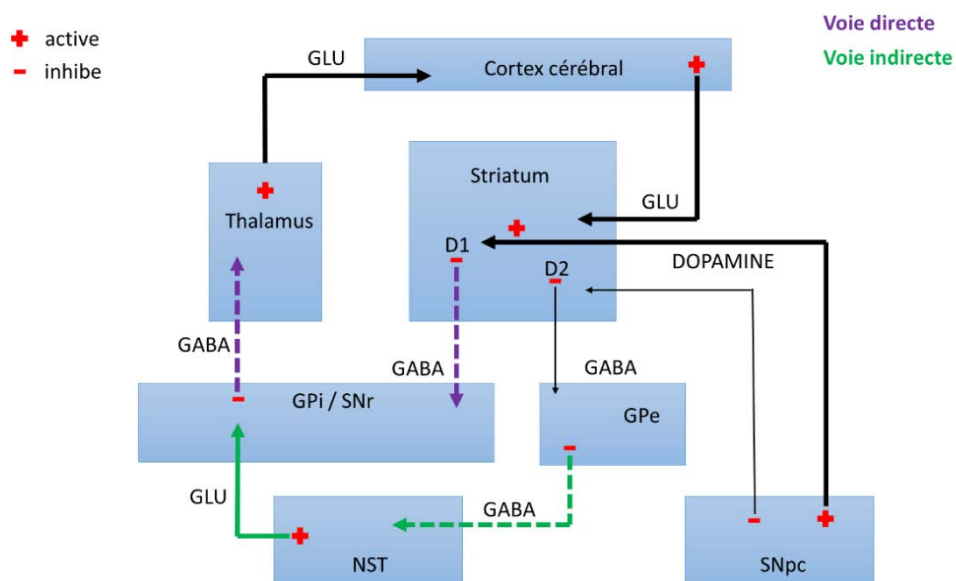
Si Edouard Brissaud suggéra dès 1895 que les lésions de la substance noire caractérisaient la MP, c'est Trétiakoff<sup>14</sup>, par la mise en évidence de la dégénérescence des neurones dopaminergiques et la présence d'agrégats protéiques dans cette région, qui apporta la preuve du lien entre la *SNpc* et la MP. La substance noire fait partie d'un groupe de noyaux sous-corticaux appelés ganglions de la base ou noyaux gris centraux (figure 1).



#### Figure 1: Organisation des noyaux gris centraux

*Les noyaux gris centraux ou ganglions de la base sont constitués d'un ensemble de structures sous-corticales qui incluent la substance noire (noir), le noyau sous-thalamique (rouge), le striatum (bleu) et le pallidum (vert). Toutes ces structures interconnectées envoient des signaux vers le thalamus (violet) qui projette dans le cortex moteur impliqué dans le contrôle de la motricité*

Le striatum comprend l'association du noyau caudé et du putamen, la substance noire et le pallidum étant quant à eux divisés en substance noire *pars compacta* (SNpc) et *reticulata* (SNr) et en pallidum interne (GPi) et externe (GPe). L'ensemble de ces structures constitue un réseau de neurones interconnectés au sein duquel s'échange des signaux neurochimiques. Les divers signaux qui transitent à travers les ganglions de la base sont appelés des neurotransmetteurs tels que le GABA, le glutamate et la dopamine qui vont permettre de contrôler la motricité. Si les ganglions de la base sont essentiellement connus pour leur rôle dans le contrôle de la motricité qui nous intéresse plus particulièrement (figure 2), ils ont aussi un rôle dans les fonctions cognitives et comportementales telles que le contrôle des émotions et la motivation.



### Figure 2: Circuit moteur des ganglions de la base

Dans les ganglions de la base, le striatum constitue la structure centrale de ce circuit moteur. Il prend naissance dans le cortex dont les neurones envoient des projections excitatrices glutamatergiques (GLU) vers le striatum. Ces informations sont ensuite relayées au thalamus moteur par l'intermédiaire de deux voies distinctes : la voie directe et la voie indirecte. Par la voie directe, les neurones gabaergiques du striatum envoient leurs projections inhibitrices directement vers la substance noire reticulata (SNr) et le pallidum interne (GPi). La voie indirecte quant à elle, connecte le striatum à la SNr et au GPi en passant par le pallidum externe (GPe) et le noyau sous-thalamique (NST). Le GPe et le NST sont deux structures connectées, le GPe envoyant un signal gabaergique (GABA) inhibiteur dans le NST qui en réponse va stimuler le GPi et la SNr via des afférences glutamatergiques excitatrices. Ces deux voies de relais de l'information cortico-striatale ont des effets antagonistes. Il a été suggéré que la voie directe agit comme un inducteur du mouvement alors que l'activation de la voie indirecte à l'effet inverse. Une fois le signal transmis au thalamus, celui-ci envoie des projections dans le cortex moteur et pré-moteur qui contrôle le mouvement.

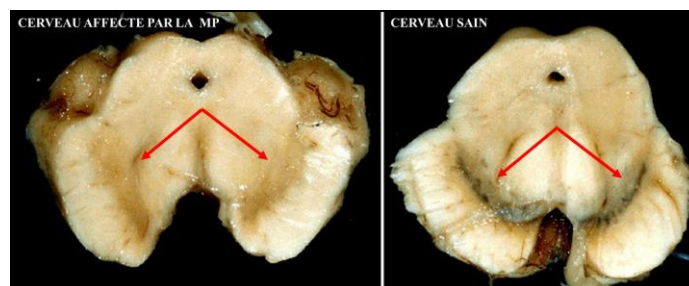
Outre les afférences provenant du cortex, le striatum reçoit également des signaux dopaminergiques provenant des neurones de la SNpc. Ces neurones capables de synthétiser la dopamine assurent le bon fonctionnement et l'équilibre entre la voie directe et indirecte. En effet,

les neurones gabaergiques du striatum expriment également des récepteurs dopaminergiques de type D1 et de type D2<sup>15</sup> qui vont moduler l'activité des ganglions de la base. Les neurones à récepteurs de type D1 sont associés à la voie directe, les neurones à récepteurs de type D2 étant eux rattachés à la voie indirecte. Ainsi, dans le striatum, la dopamine contribue à l'initiation du mouvement via les neurones à récepteurs de type D1, à l'inverse des neurones à récepteurs de type D2 qui conduisent à son inhibition. La découverte de Hornykiewicz *et al.*<sup>16</sup> fut un argument supplémentaire en faveur du rôle de la *SNpc* dans la MP ainsi que de l'importance de la dopamine puisqu'ils ont été les premiers à démontrer la baisse d'environ 80% de la dopamine dans le striatum des patients parkinsoniens, résultant de la présence des lésions au niveau de la *SNpc*<sup>17,18</sup>.

Ces lésions ont été révélées par plusieurs études neuroanatomiques de cerveaux *post-mortem* de parkinsoniens qui ont mis en évidence la mort majoritaire des neurones dopaminergiques de la *SNpc* ainsi que la présence d'agrégats protéiques nommés corps de Lewy dans les neurones persistants.

#### 1.1.2.2.2. La dégénérescence neuronale

La mort des cellules neuronales associée à la MP a été mise en évidence par Constantin Trétiakoff, qui au cours de ses travaux de thèse en 1923, avait étudié 54 cerveaux *post-mortem* incluant des cerveaux de malades parkinsoniens et des sujets sains. Ses observations lui avaient permis de montrer une différence de pigmentation entre la *SNpc* des malades et des sujets sains<sup>14,19</sup> (figure 3).



**Figure 3: Coupe transversale de cerveaux de sujets parkinsoniens et de sujets sains**

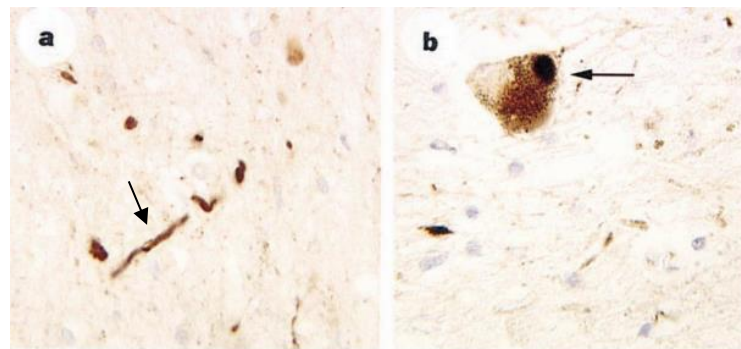
*Les flèches rouges indiquent la substance noire dépigmentée chez les parkinsoniens (à gauche) et noirâtre chez les sujets sains (à droite)*

La pigmentation de la *SNpc* résulte de la présence, dans les neurones dopaminergiques, de granules cytoplasmiques contenant un pigment noirâtre, la neuromélanine. Ainsi, la perte des neurones dopaminergiques dans la *SNpc* des malades explique son apparente dépigmentation par rapport aux

sujets sains<sup>20</sup>. Bien que cette région soit la principale affectée dans la maladie<sup>21</sup>, nous verrons par la suite qu'elle n'est pas la seule à être touchée.

#### 1.1.2.2.3. Les agrégats protéiques : corps et neurites de Lewy

Trétiakoff a aussi démontré que les inclusions (corps de Lewy) initialement décrites par Friedrich Lewy en 1912 dans le noyau basal de Meynert de cerveaux de parkinsoniens étaient aussi présentes dans la *SNpc*. Ces corps de Lewy sont des inclusions sphériques éosinophiles d'une dizaine de micromètres, entourées d'un halo clair à la périphérie (figure 4).



#### Figure 4: Agrégats protéiques caractéristiques de la maladie de Parkinson

*Les neurites de Lewy (4a) et les corps de Lewy (4b) constituent les agrégats protéiques observés à l'analyse des cerveaux post-mortem de malades atteints de la MP. Ils sont ici indiqués par les flèches noires*

Selon leur localisation dans le neurone, leur dénomination diffère : on parle de corps de Lewy lorsqu'ils sont présents dans les corps cellulaires et de neurites de Lewy lorsqu'on les retrouve dans les terminaisons axonales. Des études protéomiques et immunologiques ont révélé que ces agrégats étaient essentiellement composés de lipides et de protéines en notant que le constituant majeur est l' $\alpha$ -synucléine qui joue un rôle central dans la maladie<sup>22,23</sup>.

Si l'analyse *post-mortem* des caractéristiques anatomopathologiques de la MP conduit au diagnostic de certitude, seul l'examen clinique des atteintes motrices permet d'établir un diagnostic du vivant du patient.

#### 1.1.2.3. Le diagnostic clinique

A l'heure actuelle aucun test clinique ne permet d'établir le diagnostic de la maladie. Il ne repose donc que sur la simple observation de la tétrade symptomatologique caractéristique. Cependant, ces symptômes sont partagés par d'autres maladies neurodégénératives, regroupées

sous le terme de syndromes parkinsoniens ; la démence à corps de Lewy (DCL), l'atrophie multi-systématisée (AMS), la paralysie supranucléaire (PSP) et la dégénérescence cortico-basale (DCB) étant les plus communs. Elles peuvent néanmoins se distinguer de la MP par une évolution plus rapide des symptômes ainsi qu'une faible ou absence totale de réponse aux traitements anti-parkinsoniens. La diffusion différente des lésions explique qu'ils soient caractérisés par d'autres signes : la démence dans la DCL, des troubles cérébelleux et des signes précoces et sévères d'atteintes du système nerveux autonome dans l'AMS, des troubles oculomoteurs et des chutes précoces dans la PSP, une dysautonomie et une apraxie dans la DCB. Les lésions anatomiques sont également différentes : nombreux agrégats d' $\alpha$ -synucléine dans les neurones corticaux dans les cas de DCL ou dans les oligodendrocytes dans le cas de l'AMS. Le diagnostic de la MP nécessite alors un diagnostic différentiel pour exclure la possibilité d'être face à un syndrome parkinsonien dont les similitudes avec la MP sont à l'origine d'erreurs de diagnostic. Afin de faciliter la détection de la maladie, plusieurs études ont proposé des critères aidant à établir le diagnostic de la MP<sup>24,25,18</sup>, les plus utilisés étant les critères de l'UKPDSBB<sup>24</sup> (United Kingdom Parkinson's Disease Society Brain Bank, ANNEXE 1) qui contiennent des critères caractéristiques de la MP ainsi que des critères d'exclusions permettant d'écarter la possibilité d'un syndrome parkinsonien.

La pose du diagnostic est nécessaire à la prise en charge thérapeutique des patients et au suivi de l'évolution de la maladie. Plusieurs échelles d'évaluation de la progression de la maladie ont également été développées ; les plus utilisées étant les échelles Hoehn et Yahr<sup>26,27</sup> (ANNEXE 2) et UPDRS<sup>28</sup> (Unified Parkinson Disease Rating Scale, ANNEXE 3) qui facilitent le suivi, la mise en place et l'adaptation des traitements.

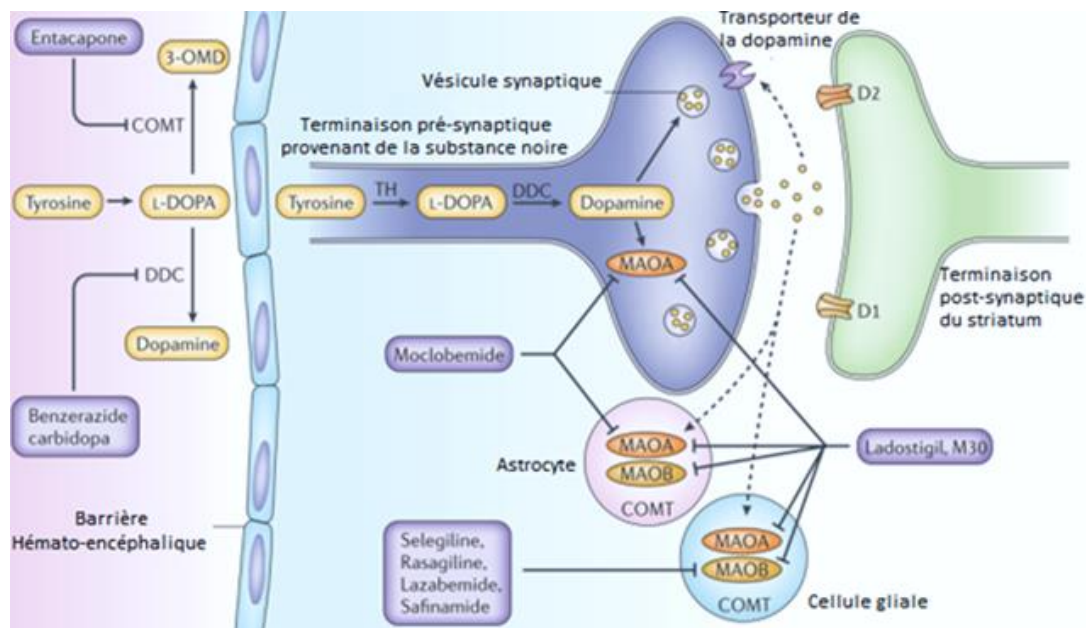
#### 1.1.2.4. Les traitements

Le traitement de la MP n'est aujourd'hui que symptomatique, reposant essentiellement sur la correction du déficit dopaminergique. Par un contrôle satisfaisant des signes moteurs, il permet de rendre aux patients et ce durant plusieurs années (la « lune de miel »), une qualité de vie satisfaisante et une augmentation de l'espérance de vie.

##### 1.1.2.4.1. Les traitements médicamenteux

Le déficit dopaminergique est traité par l'apport de levodopa (L-Dopa, le précurseur de la dopamine), l'inhibition de la dégradation enzymatique de la L-Dopa exogène ou de la dopamine endogène ou encore par l'usage d'agonistes dopaminergiques. Le traitement médicamenteux le plus efficace à l'heure actuelle est la L-Dopa. Ce précurseur, capable de traverser la barrière hémato-

encéphalique, va atteindre les neurones de la *SNpc* riches en enzymes nécessaires à sa conversion en dopamine (figure 5).



**Figure 5: Métabolisme de la dopamine et sites d'action des traitements médicamenteux (d'après Youdim, Edmondson & Keith, 2006)**

Dans les neurones dopaminergiques de la *SNpc*, le métabolisme de la dopamine commence par la conversion de la tyrosine en levodopa (L-Dopa) par la tyrosine hydroxylase (TH). La L-Dopa est ensuite prise en charge par la dopa décarboxylase (DDC) pour la convertir en dopamine. Cette dernière va ensuite être intégrée dans des vésicules pour être transportée via le transporteur à la dopamine vers les neurones dopaminergiques du striatum. Elle est aussi adressée aux cellules gliales et aux astrocytes. La dopamine peut également être dégradée par les monoamine oxydases de type A (MAOA) présentes dans les neurones, les cellules gliales et les astrocytes et/ou les monoamines oxydases de type B (MAOB) et la catéchol-O-méthyl transférase (COMT) présentes dans les astrocytes et les cellules gliales. En plus de la L-Dopa qui a la capacité de traverser la barrière hémato-encéphalique, d'autres traitements médicamenteux sont proposés aux patients et agissent comme inhibiteurs des enzymes du catabolisme de la dopamine et de la L-dopa (encadrés violets). Ils permettent de rallonger les effets de la dopamine et de la L-Dopa. Ces inhibiteurs agissent également sur les enzymes présentes dans le sang circulant en empêchant la conversion de la L-Dopa en 3-O-méthyl-dopa (3-OMD) ou en dopamine. Cette dernière qui ne peut pas passer la barrière hémato-encéphalique est alors maintenue dans la circulation sanguine et n'atteint pas ses récepteurs post-synaptiques où elle exerce sa fonction

Administrée par voie orale, elle est associée à un inhibiteur de décarboxylase périphérique pour limiter l'apparition d'effets périphériques tels que la nausée et l'hypotension ; ce qui la rend d'avantage disponible pour le système nerveux central. La L-Dopa s'est avérée bénéfique pour les patients notamment en réduisant la gravité des troubles moteurs tels que le tremblement et la rigidité<sup>29</sup>. Cependant, son efficacité s'amenuise avec le temps et s'accompagne d'effets secondaires après 4 à 5 ans de prise du traitement. Les complications motrices sont les mieux connues. Il s'agit

notamment des dyskinésies, mouvements involontaires sous-tendus soit par un excès de stimulation (cas le plus fréquent) soit par une stimulation dopaminergique insuffisante. Plus rarement la L-Dopa peut être responsable, surtout dans les cas évolués avec troubles cognitifs, de manifestations neuropsychiatriques. Les complications motrices sont d'autant plus fréquentes que le sujet est jeune et les doses fortes. En conséquence, la prise de décision quant à l'administration de ce traitement doit être réfléchie et tient compte du bénéfice apporté par rapport à la gêne qu'il peut engendrer. La L-dopa n'est donc pas toujours le traitement proposé dès les premiers signes de la maladie et est substituée par d'autres traitements tels que les agonistes dopaminergiques et les inhibiteurs des enzymes du catabolisme de la dopamine (figure 5).

Ces traitements alternatifs sont souvent les premiers traitements proposés aux patients jusqu'à ce que leur perte d'efficacité nécessite un complément par la L-dopa. Les agonistes dopaminergiques agissent à la place de la dopamine manquante au niveau des récepteurs synaptiques et sont administrés en monothérapie ou en association à de faible dose de L-Dopa pour limiter les dyskinésies. Ces traitements peuvent être à l'origine de troubles psychiques (confusion, hallucination) surtout chez le sujet âgé, qui plus est avec des troubles cognitifs. Ils exposent surtout à des troubles du sommeil (sommolence, attaques de sommeil) et plus encore à des troubles du contrôle des impulsions : appétence pour le jeu, dépenses incontrôlées, hypersexualité). Des inhibiteurs des enzymes de dégradation de la dopamine peuvent également être proposés, augmentant ainsi le potentiel et la durée d'action de la dopamine endogène (figure 5). Ces traitements alternatifs, moins efficaces que la L-dopa, sont cependant bien tolérés par les patients.

#### 1.1.2.4.2. Les traitements chirurgicaux

En complément de ces traitements médicamenteux, des méthodes thérapeutiques plus invasives ont été développées. Ceci inclut principalement la stimulation cérébrale profonde (SCP), l'apomorphine et la duodopa<sup>30</sup>. Ces traitements ne sont proposés qu'en dernier recours, aux patients à des stades avancés de la maladie ou pour lesquels plus aucune thérapie médicamenteuse ne peut être suffisamment efficace. En effet, comme tous traitements chirurgicaux, ces opérations constituent un danger et sont sources de complications. Elles sont donc effectuées de manière contrôlée et le choix des patients pouvant en bénéficier reposent sur plusieurs critères tenant compte de l'âge du patients et de la présence d'antécédents psychiatriques et de troubles cognitifs qui peuvent être rédibitoires.

La stimulation cérébrale profonde consiste en une implantation d'électrodes au sein des ganglions de la base permettant la stimulation des noyaux sous-thalamiques ou du pallidum interne.



Ceci a pour but de rétablir une fonction motrice physiologique. Ce traitement s'est avéré bénéfique notamment dans la réduction des symptômes moteurs ainsi que dans le traitement des effets secondaires aux traitements médicamenteux (dyskinésies). L'apomorphine est un agoniste dopaminergique qui peut être administrée par voie orale, par intraveineuse ou à l'aide d'une pompe sous-cutanée implantée par méthode chirurgicale<sup>31</sup>. Comme la stimulation cérébrale profonde, ce traitement permet de réduire voire d'annuler la prise de L-dopa. La duodopa est une combinaison de L-dopa et de carbidopa (un stabilisateur de la L-Dopa) qui est administrée par une endoscopie gastrotomie à l'aide d'une pompe et d'un cathéter introduit dans le duodénum. Ce traitement permet de réduire l'immobilisation des patients induite par le déficit en dopamine et améliore les symptômes non moteurs tels que les troubles du sommeil, la douleur, les troubles gastro-intestinaux.

#### 1.1.2.4.3. La vaccination

L'avancée la plus récente en matière de thérapie est l'élaboration de vaccins contre la MP par injection d'anticorps monoclonaux dirigés contre l' $\alpha$ -synucléine<sup>32</sup>. AFFITOPE PD01 est le premier vaccin élaboré dans le cadre du traitement de la MP et de l'AMS<sup>33</sup>. Des essais pré-cliniques par administration sous-cutanée d'AFFITOPE PD01 ont démontré une spécificité de ce vaccin dans l'induction d'une réponse immunitaire humorale à l' $\alpha$ -synucléine. Les tests effectués sur des souris présentant des accumulations d' $\alpha$ -synucléine montrent que ce vaccin induit une réduction de la quantité d' $\alpha$ -synucléine au niveau cérébral ainsi qu'une réduction des altérations induites par la neurotoxicité de l' $\alpha$ -synucléine telle que la mort cellulaire<sup>34</sup>. Ces tests s'avèrent donc prometteurs.

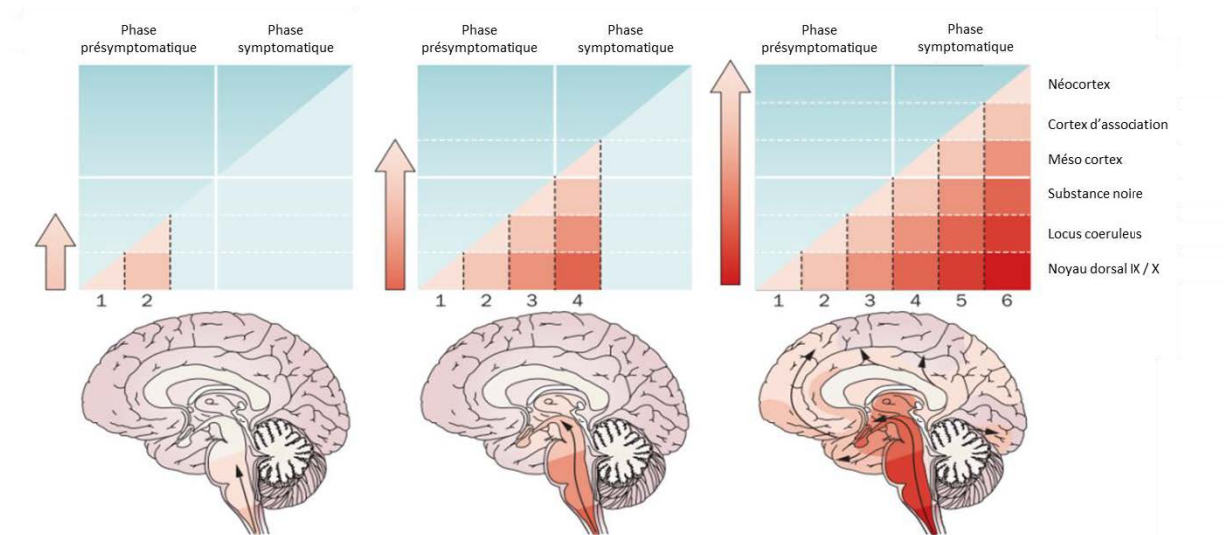
Bien que ces traitements soient bénéfiques sur le plan de la qualité de vie des patients, ils ne permettent qu'une prise en charge des symptômes. Seule la rasagiline (un inhibiteur de la MAOB) a montré des effets retardateurs de la progression de la maladie<sup>35</sup>. Cependant, le diagnostic étant posé au moment de l'apparition des signes moteurs survenant lorsque la mort neuronale a déjà atteint 50%, l'efficacité de tels traitements ne pourrait qu'être modérée. Le développement d'outils de diagnostic précoce est une nécessité afin de pouvoir maximiser l'effet des traitements neuroprotecteurs et neuromodulateurs dès lors qu'ils seront disponibles.

### 1.1.3. La maladie de Parkinson : une maladie plus complexe

#### 1.1.3.1. Les symptômes non moteurs

Si les signes non moteurs avaient déjà été décrits dès la découverte de la MP, le fait qu'ils soient décelés de façon récurrente avant l'apparition des symptômes moteurs a conduit à reconsidérer l'origine lésionnelle de la maladie. Bien que son schéma de progression soit encore sujet

à débat, Braak *et al.* ont fait le postulat d'une origine lésionnelle démarrant au niveau du système nerveux périphérique pour s'étendre à différentes aires du système nerveux central (figure 6) après y avoir observé la présence d'agrégats d' $\alpha$ -synucléine corrélant avec certains symptômes non moteurs<sup>36</sup>.



**Figure 6: Progression de la pathologie  $\alpha$ -synucléine selon l'hypothèse de Braak (d'après Goedert, Spillantini, Tredici & Heiko Braak, 2013)**

*En analysant la localisation de l' $\alpha$ -synucléine dans le cerveau, Braak a émis l'hypothèse selon laquelle les agrégats d' $\alpha$ -synucléine prendraient leur origine dans le système nerveux périphérique. Ces lésions progresseraient ensuite pour atteindre le noyau dorsal du vague et de la zone réticulée intermédiaire dans le bulbe rachidien (stade 1). Au stade 2, le locus coeruleus serait également atteints pour s'étendre au mésencéphale (dans la substance noire), du noyau tubéromammillaire de l'hypothalamus et des noyaux magnocellulaires de la base du cerveau (stade 3), du nécocortex temporal interne (stade 4), des nécocortex associatifs pariétaux polymodaux (stade 5) et des cortex prémoteurs, primaires ou associatifs unimodaux (stade 6).*

#### 1.1.3.1.1. Les troubles neuro-psychiatriques

Les patients parkinsoniens sont sujets à divers troubles neuropsychiatriques qui incluent la dépression, l'anxiété, l'apathie, les hallucinations, le délire et la démence.

La dépression est très fréquente (dans 55% des cas)<sup>37</sup>, se manifestant précocement dans plus de 27,6% des cas. Les patients parkinsoniens qui souffrent de dépression éprouvent des difficultés à exprimer leurs émotions, ce qui conduit souvent à un isolement social. L'annonce du diagnostic de MP souvent mal vécue par les patients ainsi que leur entourage contribue à cet état dépressif. Cette dépression est étroitement liée à un autre symptôme non moteur de la maladie qu'est l'anxiété. La présence des lésions dans les systèmes sérotoninergiques tels que le noyau raphé, impliqué dans la régulation de l'humeur et de la vigilance, pourrait expliquer ces troubles thymiques.

L'anxiété est deux fois plus fréquente chez les sujets parkinsoniens que dans la population générale, en particulier chez les jeunes parkinsoniens. Comme la dépression, elle est handicapante car participe à un isolement social et à des comportements agressifs. Il semblerait que les patients sujets à des phobies ou à des moments de panique constituent un terrain favorable au développement de l'anxiété.

L'apathie est de connaissance plus récente. Elle consiste en une perte de motivation qui est présente chez 30 à 40% des patients parkinsoniens. Elle est souvent confondue avec la dépression car il est difficile de distinguer l'apathie en tant que symptôme de la maladie, de la perte de motivation due à un état dépressif. Toutefois, l'état dépressif implique une détresse émotionnelle et un sentiment de mal-être, ce qui n'est pas le cas de l'apathie et permet parfois de les discriminer<sup>38</sup>.

La psychose est définie par la présence soit d'illusions (mauvaise perception des stimuli externes) soit d'hallucinations (perception sensorielle sans réalité, sans objet) ou les deux. Les hallucinations peuvent être visuelles (les plus fréquentes), auditives, tactiles ou olfactives. L'âge avancé, la durée de la maladie, les troubles du sommeil, les traitements (agonistes dopaminergiques notamment) et surtout les troubles cognitifs favorisent la survenue des hallucinations. Les altérations des fonctions cognitives comportent un ralentissement du traitement de l'information, des troubles de la mémoire, de l'attention et la capacité d'agir face à une situation nouvelle et stressante (résolution de problèmes). Ces perturbations aboutissent dans 30% des cas à une démence et peuvent être expliquées par la présence d'agrégats d' $\alpha$ -synucléine dans le locus coeruleus qui contrôle la peur et l'anxiété.

#### 1.1.3.1.2. Les troubles du sommeil

Les troubles du sommeil sont 2 à 3 fois plus présents chez les parkinsoniens par rapport à la population générale. Ils se manifestent le plus souvent au cours de la phase de sommeil paradoxal au cours de laquelle se produisent les rêves. A l'inverse d'un individu sain, le tonus musculaire des parkinsoniens n'est pas aboli au cours de cette phase de sommeil paradoxal. Les parkinsoniens sont ainsi dans un état assimilable à un état d'éveil durant lequel ils sont capables de crier, de rire, de pleurer ou de s'agiter<sup>39</sup>. De ce fait, les insomnies sont assez fréquentes (40 à 80% des cas) et se manifestent par un excès de somnolence diurne et une fatigue constante. L'origine de ces troubles n'est pas totalement connue toutefois, le locus coeruleus et le noyau raphe ont un rôle critique dans le cycle du sommeil<sup>40</sup>; la dégénérescence observée dans ces noyaux peut être à l'origine de ces troubles qui peuvent être aggravés par les traitements dopaminergiques tels que la L-dopa.

#### 1.1.3.1.3. Les troubles sensoriels

Les troubles sensoriels sont divers et incluent la douleur, les troubles visuels et olfactifs tels que l'hyposmie (forte diminution de l'odorat) et l'anosmie (perte totale de l'odorat). Les douleurs chroniques sont présentes chez 46% des patients parkinsoniens, les plus communes étant les douleurs musculo-squelettiques (crampes, courbatures) résultant d'une combinaison de facteurs tels que la rigidité et les troubles articulaires. Présents chez 90% des patients parkinsoniens, les troubles de l'odorat constituent un marqueur précoce de la maladie prédisant l'apparition des premiers symptômes dans les 4 ans à venir<sup>41</sup>. En effet, ces symptômes surviennent très précocement au cours de la maladie et semblent être associés à la présence d'agrégats d' $\alpha$ -synucléine dans les bulbes olfactifs<sup>42</sup>. Certains pensent même qu'il s'agirait du premier symptôme de la MP<sup>43</sup> du fait de la présence de ces agrégats et de la réduction du volume des bulbes olfactifs, dans le cerveau de personnes saines décédées et ayant eu des troubles de l'odorat<sup>44</sup>. Ceci est également en accord avec une hypothèse émise par certains chercheurs d'une origine virale de la MP dont le point d'entrée se ferait par voie nasale.

#### 1.1.3.1.4. Les troubles du système nerveux autonome

Ces perturbations incluent les troubles urinaires, les dysfonctions sexuelles, l'hypersudation et l'hypotension orthostatique. Les troubles urinaires se manifestent par une pollakiurie (fréquence excessive de la miction) avec parfois des fuites sinon une incontinence et plus rarement une dysurie retrouvée chez la moitié des patients. Plus de la moitié des parkinsoniens présentent des troubles du comportement sexuel allant d'une diminution de la libido à des troubles de l'érection et de l'éjaculation. Ces dysfonctions érectiles sont associées à une augmentation du risque de développer la MP chez les hommes. Enfin, l'hypotension orthostatique (baisse de la pression sanguine systolique et diastolique) est observée chez 20 à 58% des patients parkinsoniens.

#### 1.1.3.1.5. Les troubles gastro-intestinaux

Les troubles gastro-intestinaux observés chez les parkinsoniens sont essentiellement des perturbations de la motilité intestinale, l'hypersalivation, des problèmes de déglutition, des nausées et des vomissements. Ces symptômes sont liés entre eux comme la constipation qui résulte de la réduction de la motilité intestinale. Ces troubles gastro-intestinaux sont souvent associés aux troubles de l'odorat, tous deux ayant été proposés comme marqueurs précoces de la MP. Une explication à cette association a été avancée par Braak *et al.* suggérant l'implication d'un agent environnemental toxique qui entrerait dans l'organisme par voie nasale et atteindrait le système

nerveux entérique en transitant par le nerf vague. Si au niveau central, les bulbes olfactifs semblent être les premiers atteints par la pathologie  $\alpha$ -synucléine, les neurones du système nerveux entérique en sont la première cible au niveau périphérique. La possibilité d'un diagnostic précoce par biopsie du colon a donc été envisagée ; les études sont actuellement en cours pour déterminer la spécificité d'un tel diagnostic<sup>45</sup>.

#### 1.1.3.2. Les signes anatomopathologiques : la réponse inflammatoire

Bien que les premières données aient considéré la neuroinflammation comme une conséquence de la maladie, de plus en plus de données tendent à dire qu'elle contribuerait à son développement et qu'elle serait à l'origine de certains symptômes non moteurs. La présence des lymphocytes T CD4<sup>+</sup> et CD8<sup>+</sup> au niveau des cerveaux de sujets parkinsoniens suggère une migration de ces cellules immunitaires vers un site lésionnel qui corrèle avec l'hypothèse de la neuroinflammation en tant qu'acteur de physiopathologie de la MP<sup>46</sup>.

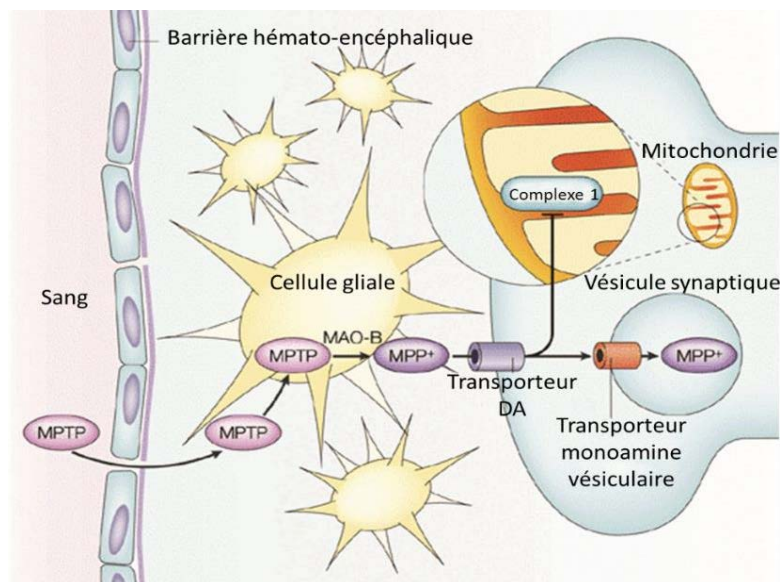
La *SNpc* des parkinsoniens est aussi marquée par la présence accrue de nombreuses cytokines proinflammatoires telles que le TNF- $\alpha$ , l'IL-1 $\beta$ , l'IFN $\gamma$  au niveau de la microglie et à l'expression de marqueurs de l'inflammation tels que la protéine GFAP (glial fibrillary acidic protein) produite par les astrocytes. Plusieurs de ces cytokines telles que le TNF- $\alpha$ , l'IL-1 $\beta$ , l'IL-6 ainsi que le facteur de transcription NF-kB (très connu pour son rôle dans l'induction de l'expression de gènes en réponse à l'inflammation)<sup>46</sup>, l'activation de la microglie et les niveaux augmentés de molécules pro-inflammatoires ont été montrés comme des activateurs de la neurotoxicité induite par l' $\alpha$ -synucléine<sup>47</sup>, manifestant un lien entre la pathologie  $\alpha$ -synucléine et l'inflammation. En effet, si Braak a fait le postulat que la pathologie  $\alpha$ -synucléine se propageait de la périphérie vers le centre, d'autres ont suggéré que la neuroinflammation se propageait selon le même schéma évolutif. Ainsi, selon les stades de Braak, les niveaux de certaines cytokines pro-inflammatoires corrèlent avec certains symptômes non moteurs ; comme les niveaux de TNF- $\alpha$  qui corrèlent avec les troubles de l'humeur et les troubles cognitifs ou les niveaux d'IL-6 plasmatique avec les troubles du sommeil<sup>48</sup>.

Aujourd'hui, si les connaissances cliniques et anatomo-pathologiques de la MP se sont grandement améliorées, les connaissances des mécanismes physiopathologiques sont encore incomplètes. Néanmoins, la découverte des déterminants environnementaux et génétiques a joué un rôle important et a grandement contribué aux connaissances actuelles de ces mécanismes.

#### 1.1.4. Les mécanismes moléculaires contribuant au développement de la MP

##### 1.1.4.1. L'apport des facteurs environnementaux

La MP a longtemps été considérée comme la résultante d'une exposition à un ou plusieurs facteurs environnementaux. Malgré des différences cliniques et neuropathologiques leur rôle a en effet été évoqué devant l'apparition de syndromes parkinsoniens chez des sujets exposés au manganèse<sup>49,50,51,52</sup> ou ayant présenté une encéphalite de Von Economo<sup>53</sup>. Après les études épidémiologiques de Barbeau démontrant qu'au Québec l'incidence variait selon les bassins hydrographiques, c'est l'observation de cas de MP survenant suite à la consommation de MPTP qui a été l'argument majeur de la contribution de l'environnement à la genèse de la MP<sup>54,55</sup>. En effet, en 1982 ont été rapportés des cas de MP chez des étudiants californiens consommateurs d'héroïne, qui après avoir tenté de synthétiser leur propre drogue avaient produit puis utilisé le MPTP (1-méthyl-4-phényl-1,2,3,6-tétrahydropyridine) quelques jours avant de présenter les symptômes moteurs caractéristiques de la maladie. A la suite de cette découverte, de nombreuses recherches ont été menées afin de comprendre le mécanisme par lequel le MPTP était capable d'induire une MP. Une fois dans l'organisme, celui-ci est capable d'atteindre le cerveau via la circulation sanguine. Arrivé au cerveau, il va être pris en charge par les cellules gliales de la *SNpc* contenant la monoamine oxydase de type B (MAOB). Cette enzyme catalyse la conversion du MPTP en une molécule toxique, le MPP<sup>+</sup> (figure 7).

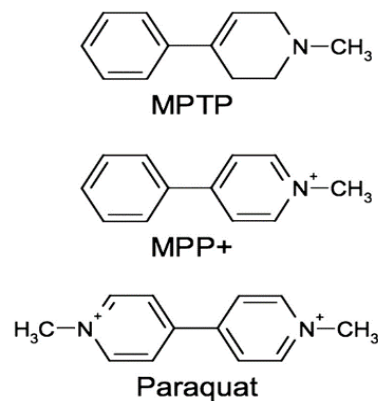


**Figure 7: Synthèse et acheminement du MPP<sup>+</sup> vers les neurones dopaminergiques de la *SNpc* (d'après Vila & Przedborski, 2003)**

*Le MPP<sup>+</sup> est synthétisé dans les cellules gliales par conversion du MPTP en MPP<sup>+</sup> par l'enzyme MAOB (monoamine oxydase de type B). Une fois synthétisé, il va être acheminé vers les neurones*

dopaminergiques en passant à travers le transporteur membranaire à la dopamine (DA) situé à la surface de ces neurones.

Dans les neurones, le MPP<sup>+</sup> altère la chaîne respiratoire mitochondriale par inhibition du complexe 1. La caractérisation de la structure et du mode d'action du MPTP a permis de mettre en lumière le rôle potentiel des perturbations de la mitochondrie mais également d'identifier d'autres molécules capables d'induire les symptômes de la maladie. En effet, d'autres molécules de structure chimique similaire (figure 8) ou ayant également une action inhibitrice du complexe 1 de la mitochondrie sont capables de mimer la MP.

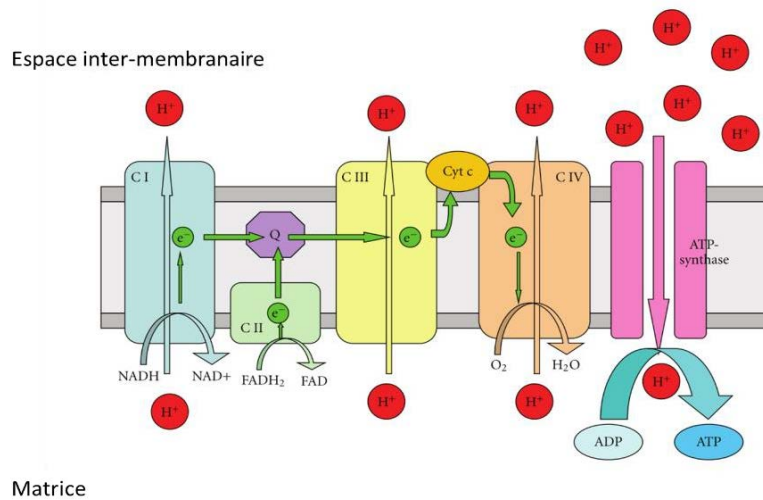


**Figure 8: Similarité de structure chimique entre le MPTP, le MPP<sup>+</sup> et le paraquat**

C'est le cas en particulier de certains produits utilisés en agriculture industrielle comme le paraquat (herbicide), la roténone (pesticide naturel) ou encore le maneb (fongicide) pour lesquels des études épidémiologiques notamment au sein de populations d'agriculteurs, ont montré que leur exposition était associée à un risque augmenté de MP. L'implication de solvants industriels (trichloréthylène et perchloroéthylène) ainsi que de gaz (monoxyde d'azote) a été énoncée mais n'est pas encore démontrée.

La mitochondrie a de nombreux rôles dans la cellule. Cette organelle intracellulaire est constituée de deux membranes phospholipidiques, la membrane interne et la membrane externe et est impliquée dans la production d'énergie (production d'ATP par phosphorylation oxydative), la régulation de la mort cellulaire, le métabolisme calcique et la production d'espèce réactive de l'oxygène (ROS) et de l'azote (RNS)<sup>56</sup>. La phosphorylation oxydative est la principale fonction de la mitochondrie et permet la synthèse d'ATP, principale source d'énergie de la cellule. Cette synthèse fait appel à 5 complexes protéiques différents, localisés dans la membrane interne de la mitochondrie. Ils forment la chaîne respiratoire mitochondriale, le long de laquelle s'effectue un

transfert d'électrons associé à un échange de protons  $H^+$  entre la matrice mitochondriale et l'espace intermembranaire. Cet échange est utile à l'ATP-synthase pour la production d'ATP<sup>57</sup> (figure 9).



**Figure 9: Biosynthèse de l'ATP à travers la chaîne respiratoire mitochondriale (d'après Keane, Kurzawa, Blain, & Morris, 2011)**

La synthèse d'ATP à travers la chaîne respiratoire mitochondriale fait intervenir un transport d'électrons à travers les différents complexes de la membrane interne de la mitochondrie. Le transport d'électrons à travers le complexe 1 (C I) et le complexe 2 (C II) est généré par la conversion du  $NADH$  en  $NAD^+$  par la NADH ubiquinone reductase (complexe 1) et par la conversion du  $FADH_2$  et  $FAD$  via la succinate ubiquinone reductase (complexe 2) et l'intervention du coenzyme Q. Ces électrons sont ensuite transférés vers le cytochrome c par l'enzyme ubiquinol cytochrome c reductase du complexe 3 (C III) pour enfin atteindre le complexe 4 (C IV) où l'enzyme cytochrome c oxydase les utilise pour convertir le dioxygène ( $O_2$ ) en eau ( $H_2O$ ). Au cours de ce transfert d'électrons, des ions  $H^+$  sont libérés à travers les complexes 1, 2 et 4 pour permettre la synthèse d'ATP par l'ATP synthase (complexe 5)

Au cours du processus de phosphorylation oxydative, des électrons peuvent s'échapper de la mitochondrie via les complexes 1 et/ou 3 et interagir avec l'oxygène et l'azote circulant, conduisant à la formation de ROS et de RNS qui en forte quantité sont toxiques pour la cellule. Dans des conditions physiologiques, la cellule met en place des mécanismes antioxydants capables de prendre en charge ces ROS et RNS pour les éliminer. Ces mécanismes de défense font intervenir des enzymes telles que la Cu/Zn-superoxyde dismutase, la Mn-superoxyde dismutase, la glutathion peroxydase, la catalase mais également des molécules telles que le glutathion, la vitamine C et la vitamine E<sup>58</sup>. Ainsi, en situation physiologique, la production d'espèces actives et les mécanismes de détoxification sont à l'équilibre pour assurer l'homéostasie de la cellule<sup>58</sup>. Un déséquilibre entre ces mécanismes de production et d'élimination est à l'origine d'un stress oxydant toxique pour la cellule.



La découverte de l'induction d'un phénotype parkinsonien par les inhibiteurs du complexe 1 de la chaîne respiratoire (MPTP, paraquat, roténone, maneb) a permis d'asseoir l'hypothèse d'un lien entre la MP et les dysfonctionnements de la mitochondrie<sup>59</sup>. L'identification de ces inhibiteurs a conduit à la création de modèles expérimentaux animaux et cellulaires qui ont aidé à mieux comprendre les mécanismes physiopathologiques en démontrant la présence d'altérations également observées chez les patients. En effet, une réduction de l'activité des protéines du complexe 1 mitochondrial à hauteur de 42% a été observée dans la *SNpc* de cerveaux de malades décédés de MP. Cette diminution peut s'accompagner d'une réduction de l'expression de plusieurs sous-unités du complexe, contribuant à l'inhibition de la respiration mitochondriale et à une production augmentée de ROS<sup>60</sup> à l'origine du stress oxydant montrant ainsi le lien entre la mitochondrie, le stress oxydant et la maladie. L'importance de ce stress est confirmée par la présence accrue de molécules anti-oxydantes telles que le glutathion ou la surexpression d'enzymes de détoxification telles que la cyclo-oxygénase COX2 ou la superoxyde dismutase SOD2 dans les cerveaux de rats ainsi que les lignées cellulaires dopaminergiques de neuroblastomes exposées au MPTP<sup>61,62</sup>, au paraquat<sup>63</sup> et à la 6-hydroxy-dopamine. En effet, la dopamine elle-même est à l'origine d'un stress oxydant au cours duquel elle est convertie en dopamine quinone et 6-hydroxy dopamine (6-OHDA) par oxydation et hydroxylation respectivement. Ces deux métabolites accentuent un stress déjà présent. Ainsi, la 6-OHDA est très utilisée dans les analyses expérimentales pour générer des modèles de la MP ainsi que la dopamine quinone, toutes deux associées à la pathogenèse de la MP du fait de leur capacité à générer des ROS<sup>64, 65,66</sup>. Le fer, très fortement concentré au niveau de la substance noire contribue à renforcer ce stress oxydant non seulement par son implication dans la synthèse de 6-OHDA<sup>67</sup> mais également via son implication dans la réaction de Fenton (réaction d'oxydo-réduction du peroxyde d'hydrogène, H<sub>2</sub>O<sub>2</sub>) qui aboutit à la synthèse de radicaux libres. Le stress oxydant induit par l'exposition à la dopamine quinone conduit à la perte de plus de 50% de l'activité de SOD2 qui corrèle avec la formation concomitante d'agrégats protéiques<sup>66</sup> caractéristiques de la maladie. Une autre conséquence de ce stress oxydant est son effet sur l'agrégation des protéines : il contribue à l'agrégation de l' $\alpha$ -synucléine et à la formation des corps de Lewy par intervention du cytochrome c dans la nitrosylation de résidus tyrosines de l' $\alpha$ -synucléine qui favorise son agrégation<sup>68,68</sup>. Toutes ces caractéristiques de la maladie induites par les troubles de la mitochondrie et le stress oxydant ont révélé leur importance dans la physiopathologie de la MP.

Si l'apport des facteurs environnementaux à une meilleure connaissance des mécanismes physiopathologiques de la MP est indéniable, ils ne peuvent cependant pas tout expliquer ; la plupart des cas observés se déclarant en l'absence de toute exposition. D'autres éléments doivent alors intervenir; c'est le cas des facteurs génétiques.

#### 1.1.4.2. L'apport des facteurs génétiques

D'abord totalement récusé, le rôle de la génétique dans le développement de la MP a émergé à la fin des années 1980. Certains chercheurs avaient déjà observé qu'environ 15% des malades avaient des antécédents connus de MP au sein de leur famille, chiffre supérieur à celui constaté dans la population générale. L'attention portée aux cas familiaux a souligné que la transmission de la maladie se faisait tantôt selon un mode autosomique dominant (le plus fréquent) tantôt selon le mode autosomique récessif. Les premières études basées sur l'analyse de plusieurs dizaines de paires de jumeaux ne se sont pas pour autant révélées convaincantes sans doute en raison de biais d'analyse identifiés par la suite<sup>69-73</sup>. C'est le groupe de Duvoisin, pourtant convaincu quelques années plus tôt que l'hérédité ne jouait aucun rôle dans la maladie, qui en analysant la famille de Contursi en 1996 a permis de mettre en évidence l'année suivante la première mutation causale de la MP<sup>74</sup>. Cette mutation dans le gène *SNCA* a marqué un tournant dans l'histoire de la MP et a été le point de départ d'une succession d'études génétiques. Par des analyses de liaison et des stratégies de gènes candidats, dix-huit loci liés à la MP nommés PARK 1 à PARK 18<sup>75</sup> ont été identifiés. Ils sont récapitulés dans le tableau ci-dessous (Tableau 1). L'analyse approfondie de ces régions génomiques a mené à la caractérisation des mutations de gènes contenus dans ces loci définissant ainsi les formes monogéniques de la maladie.

Nom	Locus	Gène	Types de mutation	Transmission	Phénotype clinique
PARK1	4q22.1	SNCA	mutations A30P, E46K, A53T	AD	MP ou démence à corps de Lewy
PARK2	6q26	PRKN	>100 mutations (mutations ponctuelles, réarrangements d'exons)	AR	MP de début précoce
PARK3	2p13	??		AD	MP
PARK4	4q22.1	SNCA	duplications, triplications	AD	MP, démence corrélée au dosage génique
PARK5	4p13	UCHL1	substitution I93M	AD?	MP
PARK6	1p36.12	PINK1	>20 mutations (mutations ponctuelles, réarrangements d'exons)	AR	MP de début précoce
PARK7	1p36.23	DJ-1	mutations ponctuelles, délétions	AR	MP de début précoce
PARK8	12p12	LRRK2	mutations ponctuelles, G2019S la plus fréquente	AD	MP
PARK9	1p36.13	ATP13A2	mutations ponctuelles	AR	Syndrome de Kufor-Rakeb
PARK10	1p32	??		AD?	MP
PARK11	2q37.1	GIGYF2	mutations ponctuelles	AD	MP
PARK12	Xq21-q25	??		lié à l'X	MP
PARK13	2p13.1	OMI/HTRA2	mutations ponctuelles	AD?	MP
PARK14	22q13.1	PLA2G6	mutations ponctuelles	AR	dystonie, SP, pyramidal, démence
PARK15	22q12.3	FBXO7	mutations ponctuelles	AR	syndrome pyramidal et SP
PARK16	1q32	NUCKS1??		susceptibilité	MP
PARK17	16q11.2	VPS35	mutations ponctuelles	AD	MP
PARK18	3q27.1	EIF4G1 ?	mutations ponctuelles	AD	MP
	17q21.31	MAPT	mutations ponctuelles	AD	démence fronto-temporale avec SP
	15q26.1	POLG	mutations ponctuelles	AR?, AD?	SP avec neuropathie, ophtalmoplégie
	2p13	DCTN1	mutations ponctuelles	AD	Syndrome de Perry
	2q24.1	NR4A2	mutations ponctuelles	AD?	MP
	5q23.2	SNCAIP	mutations ponctuelles	??	MP
	12q24.12	ATXN2	expansion de triplets CAG ?	AD	ataxie spino-cérébelleuse, MP
	14q32.12	ATXN3	expansion de triplets CAG ?	AD	ataxie spino-cérébelleuse, MP
	1q22	GBA	mutations ponctuelles	??	maladie de Gaucher, MP

**Tableau 1: Loci PARK liés à la maladie de Parkinson d'après le recueil OMIM (Online Mendelian Inheritance in Man, <http://omim.org/>)**

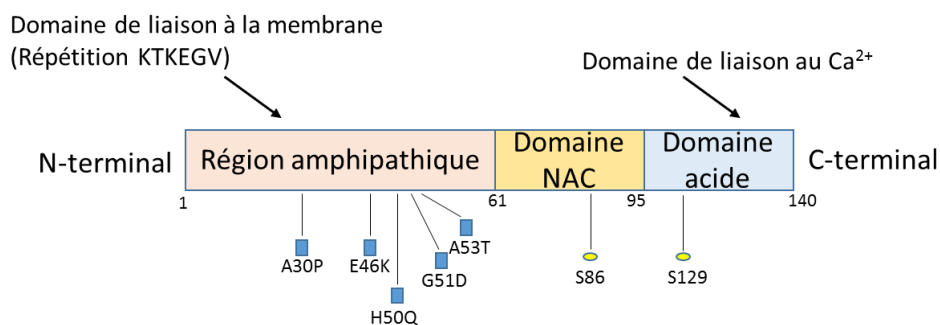
*Les gènes des 18 loci liés aux formes autosomiques dominantes (AD) et récessives (AR) ainsi que certains gènes de susceptibilité de la maladie sont indiqués dans le tableau ci-dessus. Les types de mutations et les phénotypes cliniques associés y sont également répertoriés.*

Ces mutations se distinguent par leur caractère autosomique dominant (mutation de *SNCA*, *LRRK2*, ...) ou récessif (*PRKN*, *PINK1*, *DJ-1*...) et par l'induction d'un phénotype purement parkinsonien

ou tantôt évocateur d'une maladie apparentée. Ainsi, bien que plusieurs gènes aient été liés à la MP seules les mutations des gènes *SNCA*, *LRRK2*, *PRKN*, *PINK1*, *DJ-1* et *VPS35* ont formellement été identifiées comme mutations causales de la MP et seront discutés dans ce manuscrit. Nous aborderons plus en détail les gènes étudiés au cours des analyses de transcriptome de ce travail de thèse, à savoir les gènes *SNCA*, *LRRK2* et *ATXN2* et nous ferons une brève description de *EIF4G1* compte tenu de son importance pour le laboratoire.

#### 1.1.4.2.1. Le gène *SNCA*

*SNCA* (synuclein alpha) est le premier gène au sein duquel des mutations causales de la MP ont été identifiées. Son implication dans la maladie a été révélée par Polymeropoulos *et al.* qui, en 1997, ont identifié au sein d'une famille italienne la première mutation à l'origine de la MP<sup>74</sup>. Cette mutation, mise en évidence par liaison génétique, est une mutation faux-sens à transmission autosomique dominante qui consiste en un remplacement d'une guanine par une adénine au codon 209 (G209A). Cette mutation se traduit au niveau protéique par un changement de l'alanine en position 53 par une thréonine d'où l'appellation mutation A53T. Retrouvée chez des parkinsoniens au sein d'autres familles, cette mutation est absente chez les sujets témoins indiquant que la mutation est associée au phénotype de la MP bien que sa pénétrance ne soit pas totale (85% à 65 ans)<sup>76</sup>. A la suite de cette première découverte, d'autres mutations causales de *SNCA* ont été identifiées incluant les mutations A30P, E46K, G51D et H50Q<sup>77</sup> (figure 10).



**Figure 10: Localisation des mutations et des domaines protéiques de l'alpha-synucléine (d'après Emanuele & Chiergatti, 2015)**

*L'alpha-synucléine est caractérisée par la présence de 3 domaines fonctionnels, l'ensemble des mutations causales de la MP se situant dans le domaine d'interaction de la protéine avec les membranes. Elle contient également des acides aminés sérines (S87 et S129) phosphorylables qui ont des fonctions différentes : la phosphorylation de la sérine 129 augmentant la capacité d'agrégation de l'alpha-synucléine à l'inverse de la phosphorylation de la sérine 87*

Notre équipe et celle de Singleton avons montré que la surexpression de *SNCA* étaient responsable de la MP. En effet, des duplications<sup>78</sup> et triplications<sup>79</sup> de son locus sont liées à certaines formes génétiques et sont responsables de l'apparition de la MP. Ce gène s'étend sur une portion chromosomique de 117 kb et est localisé sur le chromosome 4q21-23<sup>80,81</sup>. Constitué de 7 exons<sup>80</sup>, son expression dans le cerveau conduit à la synthèse d'un transcrit majeur faisant 3,6kb<sup>82</sup>. Initialement décrit comme s'exprimant principalement dans le système nerveux<sup>83</sup>, l'expression de *SNCA* est également observée dans les cellules sanguines (plasma, plaquettes, globules rouges et cellules mononucléées) et les cellules épidermiques<sup>84</sup>. Son expression est modulée par plusieurs mécanismes faisant intervenir des polymorphismes génétiques en 5' et 3' du gène, des facteurs de transcription ainsi que des facteurs épigénétiques (méthylation, miARNs) et environnementaux (alcool). Le polymorphisme Rep1, correspondant à une succession de répétitions dinucléotidiques à environ 10kb en amont du site d'initiation de la traduction, est associé soit à un risque accru de développer la MP lorsque la région génotypée contenant les répétitions atteint 263pb ou un effet protecteur lorsqu'elle est de 259pb<sup>85,86,87</sup>. Rep1 n'est pas le seul polymorphisme permettant de moduler l'expression de *SNCA*. Il existe des polymorphismes nucléotidiques dans la région 3'UTR (région non traduite de l'ARNm) de *SNCA*<sup>88,89</sup>. Les études d'associations ont également identifié de nombreux polymorphismes dont certains sont capables de moduler l'expression de *SNCA*. La méthylation au niveau des îlots CpG du promoteur de *SNCA* ou dans l'intron 1 par la méthyltransférase DNMT1 (DNA (Cytosine-5-)-Methyltransferase 1) induit l'expression de *SNCA*<sup>90,91,92,93,94</sup>. Enfin, des facteurs *trans* régulateurs contrôlent également l'expressions de ce gène ; il s'agit des facteurs de transcription GATA 1 (GATA binding protein 1) et GATA 2 (GATA binding protein 2)<sup>84,95</sup> ainsi que plusieurs micro-ARNs (miARNs) tels que miR-153 et miR-7<sup>96,97,98</sup>.

L'intérêt pour *SNCA* dans l'induction de la MP s'est avéré d'autant plus important que ce gène code le constituant majeur des corps de Lewy qui sont retrouvés à la fois chez la majorité des sujets présentant des formes monogéniques et chez les sujets sporadiques représentant ainsi une caractéristique pathologique commune. L' $\alpha$ -synucléine existe sous 4 isoformes protéiques différentes générées par épissage alternatif : l'isoforme de 140kDa majoritaire au niveau cérébral et les isoformes de 126, 112 et 98kDa par exclusion soit de l'exon 3, de l'exon 5 ou les deux respectivement. Tout comme son gène, l' $\alpha$ -synucléine est exprimée à la fois au niveau cérébral où elle représente 1% des protéines cytosoliques totales et dans les cellules sanguines (érythrocytes et plaquettes). Elle est caractérisée par la présence d'au moins 3 domaines fonctionnels : un motif apolipoprotéine dans la région N-terminale, un domaine non amyloïde central et un domaine acide dans la région C-terminale (figure 10). L'extrémité N-terminale s'étend du premier au soixante et unième acide aminé et se caractérise par la présence de 7 motifs répétés de 11 acides aminés dont le

cœur est défini par un consensus KTKEGV et semble être impliqué dans la capacité de la protéine à interagir avec les lipides membranaires. L'extrémité C-terminale quant à elle est caractérisée par une haute charge négative, ces deux extrémités étant séparées par la présence d'une région centrale hydrophobe du 61<sup>ème</sup> au 95<sup>ème</sup> acide aminé : le domaine NAC (non-Ab component)<sup>81,99</sup>. Ce dernier est essentiel à l'agrégation et à la formation des corps de Lewy qui s'initie selon un processus de nucléation<sup>100</sup> favorisé par la surexpression du gène *SNCA* ainsi que par les mutations A30P et A53T<sup>101</sup>, l'augmentation du cytochrome c et du stress oxydant<sup>102</sup> et les modifications post-traductionnelles telles que la phosphorylation de la serine 129, la nitration ou encore la troncation de l'extrémité C-terminale de la protéine. Bien que l' $\alpha$ -synucléine ait longtemps été considérée comme une protéine sans conformation native, son extrémité N-terminale adopte une structure en hélice alpha qui lui confère également sa capacité à se lier à la membrane et favorise son agrégation.

Toutes les modifications génétiques causales de *SNCA* ne sont pas strictement associées à une MP. Seule la duplication de son locus et la mutation A30P conduisent à un phénotype clinique similaire aux formes sporadiques, à l'exception de l'âge d'apparition des premiers symptômes, plus précoce pour les sujets porteurs d'une duplication<sup>103,104</sup>. Ces sujets présentent aussi une dépigmentation de la *SNpc* ainsi que la présence de corps de Lewy dans la *SNpc* et autres régions apparentées à la maladie. Les sujets porteurs d'une mutation E46K ont un phénotype plus proche de la DCL de même que les sujets porteurs d'une triplication qui ont également un âge d'apparition des premiers symptômes précoces. Toutefois, l'étude de ce gène, des perturbations induites par ses mutations et de la localisation de la protéine a permis de mieux comprendre les mécanismes pouvant être impliqués dans la MP en éclairant sur les processus cellulaires dans lesquels l' $\alpha$ -synucléine est impliquée.

La première fonction attribuée à l' $\alpha$ -synucléine a été basée sur sa localisation dans la cellule. Retrouvée au niveau des terminaisons pré-synaptiques, elle a été assimilée à un rôle dans la transmission synaptique dopaminergique. Ce rôle a notamment été soutenu d'une part par l'inhibition de l'expression de *SNCA* qui a montré une inhibition du pool et de la taille des vésicules synaptiques dans les neurones hippocampiques<sup>105</sup>, d'autre part par la perte du gène chez des souris qui induit un déficit en dopamine striatale<sup>106</sup> et enfin par la capacité de l' $\alpha$ -synucléine à se lier à la tyrosine hydroxylase, enzyme limitante de la voie de biosynthèse de la dopamine. Cette liaison induite par la surexpression de la forme sauvage ou de la forme mutée A53T inhibe l'activité de la protéine induisant une réduction de la synthèse de la dopamine. L' $\alpha$ -synucléine a également la capacité de se lier au promoteur de la tyrosine hydroxylase pour réguler son expression sachant que cette enzyme a été montrée comme étant diminuée dans le cerveau de MP. L' $\alpha$ -synucléine contribue à l'assemblage du complexe formé par les protéines t-SNARE et v-SNARE qui sont essentielles à la

fusion des vésicules synaptiques<sup>107,108</sup>. Cette capacité de fusion des vésicule est inhibée par l'agrégation de l' $\alpha$ -synucléine<sup>109</sup>. Localisée dans ces vésicules, l' $\alpha$ -synucléine interagit avec les membranes phospholipidiques via son domaine apolipoprotéine<sup>110</sup>. Cette liaison aux membranes et aux vésicules est abolie par la mutation A30P<sup>111</sup> corroborant ce rôle de l' $\alpha$ -synucléine dans le trafic vésiculaire et suggérant l'implication du transport vésiculaire dans la maladie.

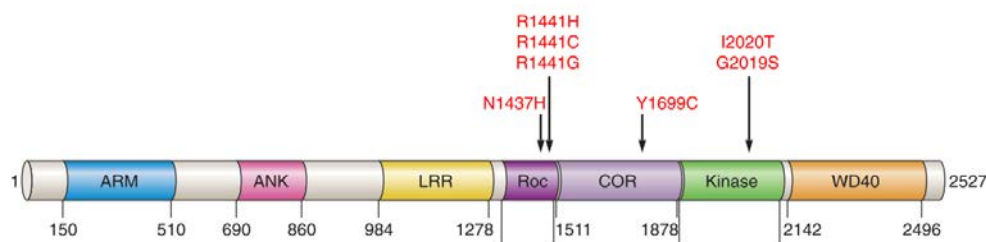
L' $\alpha$ -synucléine semble également avoir un rôle dans les mécanismes de dégradation protéique ainsi que dans la mitochondrie. En effet, la surexpression de la forme mutée A30P dans des lignées cellulaires PC12 dérivées de phéochromocytome de rat a montré une inhibition de l'activité du protéasome et une augmentation de la sensibilité des cellules à l'apoptose<sup>112</sup>. L' $\alpha$ -synucléine a été détectée dans le lumen des lysosomes (organite cellulaire impliqué dans la dégradation protéique) traduisant le recrutement de l' $\alpha$ -synucléine dans le lysosome par interaction avec les protéines Hsc70 et LAMP-2A. Ces deux protéines ont des niveaux réduits dans la SN de patients parkinsoniens<sup>113</sup> favorisant l'accumulation d' $\alpha$ -synucléine. Elle co-localise également avec LAMP-2A dans les neurones de souris exposées au paraquat ou surexprimant l' $\alpha$ -synucléine<sup>114</sup>. Les traitements de cellules avec des inhibiteurs de la mitochondrie tels que la roténone ou le paraquat induit une augmentation de l'expression de l' $\alpha$ -synucléine, augmentation qui est accompagné d'une augmentation de son agrégation dans les neurones de la *SNpc* de souris exposées à ces inhibiteurs<sup>115</sup>. Ces inclusions s'amenuisent avec le retour à une fonction mitochondriale basale marquant le lien entre mitochondrie et pathologie de l' $\alpha$ -synucléine. La surexpression de l' $\alpha$ -synucléine dans des souris invalidées pour le gène PGC-1 $\alpha$  a montré une plus forte sensibilité à la toxicité induite par l' $\alpha$ -synucléine chez ces souris (PGC-1 $\alpha$  étant une protéine impliquée entre autre dans le contrôle de la morphologie mitochondriale) et est restaurée par la surexpression de PGC1- $\alpha$ <sup>116</sup> accentuant le lien entre l' $\alpha$ -synucléine et la mitochondrie.

Enfin l'  $\alpha$ -synucléine joue un rôle majeur dans les synucléinopathies par sa capacité à se propager de cellule en cellule et de propager la pathologie  $\alpha$ -synucléine comme l'a proposé Braak. Cette propagation semble s'effectuer selon un mécanisme de types prions puisque des expériences de transplantations ont montré que les cellules saines étaient capables de présenter une agrégation initialement absente.

Ainsi bien que l' $\alpha$ -synucléine soit souvent indiquée comme une protéine dont la fonction n'est pas complètement élucidée, elle intervient dans le trafic vésiculaire, les mécanismes de dégradation protéique, en particulier le lysosome et l'autophagie et est impliquée dans les dysfonctions de la mitochondrie. De plus, elle joue un rôle primordial dans le développement et la propagation des lésions de la MP.

#### 1.1.4.2.2. Le gène *LRRK2*

*LRRK2* est le second gène porteur de mutations à transmission autosomique dominante ayant été lié à la MP. En 2004, grâce à des analyses de liaisons génétiques, deux équipes ont mis en évidence un changement de la guanine, au codon 6055, en une adénine chez des sujets atteints de la MP<sup>117,118</sup>. Cette modification au niveau du génome conduit au niveau protéique à un remplacement de l'acide aminé glycine en une serine en position 2019 (mutation G2019S). Bien qu'identifiée en premier lieu chez les malades présentant une forme génétique, cette mutation est aussi présente chez certains sujets sporadiques ainsi que chez des sujets sains. La pénétrance de la mutation varie avec l'âge et l'origine ethnique de la population considérée (de 17% à l'âge de 50 ans à 85% à l'âge de 70 ans), cela pourrait expliquer l'absence de phénotype parkinsonien chez certains sujets porteurs de la mutation<sup>119</sup>. Aujourd'hui on dénombre plus de 100 polymorphismes et/ou mutations (faux-sens ou non-sens) dans ce gène dont seules six mutations ont un rôle biologique démontré dans la MP : R1441C, R1441G, R1441H, Y1699C, G2019S et I2020T<sup>120</sup> (figure 11).



**Figure 11: Représentation de la protéine *LRRK2* et des mutations causales de la maladie de Parkinson (d'après Corti, Lesage, & Brice, 2011)**

La mutation G2019S compte pour 5-6% des formes familiales européennes et 1-2% des formes sporadiques. Elle est présente à hauteur de 40 à 50% chez les malades d'origine maghrébine ou juive ashkénaze alors qu'elle est quasiment absente dans les populations asiatiques<sup>121</sup>. Si la majorité des études indiquent que le phénotype induit par la mutation G2019S est similaire à celui des parkinsoniens sporadiques non mutés (âge d'apparition, sensibilité à la dopa-thérapie, phénotype clinique à l'exception de certains symptômes moteurs et de l'absence de corps de Lewy dans de rares cas)<sup>122</sup>, il a été rapporté que les patients porteurs d'autres mutations de *LRRK2* présentaient en plus d'un phénotype parkinsonien, des phénotypes de type paralysie supranucléaire progressive ou atrophie musculaire<sup>121</sup>.

Le gène *LRRK2* se situe sur le chromosome 12q12 et possède 51 exons codant un transcrit de 9kb<sup>123</sup>. Exprimé de manière ubiquitaire, son expression est fortement présente dans le cerveau au



niveau de la substance noire, des ganglions de la base, du cortex, de l'hippocampe et du cervelet. Les facteurs contribuant à sa régulation sont très peu connus, toutefois, Hyun Hin Cho *et al.* ont mis en évidence une régulation négative de l'expression de *LRRK2* par l'intermédiaire de miARNs dont le miR-205<sup>124</sup>. *LRRK2* code une protéine de 2527 acides aminés nommée LRRK2 ou dardarine (du basque dardara qui signifie tremblement). Elle appartient à la famille de protéines ROCO qui est caractérisée par la présence de 5 domaines conservés : LRR (leucin rich repeat), ROC (Ras in complex proteins), COR (C-terminal domain of Roc), MAPKKK (mitogen-activated protein kinase kinase kinase) et WD40 contenant des répétitions dipeptidiques de tryptophanes et d'acides aspartiques<sup>125</sup> (figure 11).

L'identification des rôles physiologiques de LRRK2 a été basée dans un premier temps sur l'analyse des domaines fonctionnels que contient la protéine. Mata *et al.* ont suggéré que la présence des domaines ARM, ANK, LRR et WD40 conférerait à la protéine la capacité d'interagir avec ses partenaires. Cette interaction lui permettrait notamment d'exercer une fonction de protéine échafaudage contribuant à la formation de complexes multiprotéiques<sup>126</sup>. Le domaine WD40 a été montré comme étant nécessaire à l'interaction de LRRK2 avec de nombreuses protéines pré-synaptiques telles que SNAP-25, syntaxine 1A, VAMP et actine<sup>127,128</sup>. Ces protéines sont impliquées dans le transport des vésicules le long du cytosquelette suggérant un rôle de LRRK2 dans le trafic des vésicules. Ce trafic permet la communication entre les différents systèmes endomembranaires de la cellule dont le réticulum endoplasmique, l'appareil de golgi et le lysosome ainsi que l'adressage des protéines dans ces différents sous-compartiments cellulaires. De nombreuses données de la littérature viennent étayer l'implication de LRRK2 dans le transport des vésicules synaptiques comme par exemple l'inhibition de LRRK2 dans un modèle *C. elegans* qui est responsable d'une perte de l'adressage des protéines vers la synapse indiquant l'implication de LRRK2 dans ce flux de protéines. Cette perturbation du trafic vésiculaire induit par l'inhibition de l'expression de LRRK2 est également observée suite à l'inhibition de son activité kinase qui empêche l'interaction de LRRK2 avec les vésicules et des protéines présynaptiques<sup>127</sup>. Cet adressage semble faire intervenir l'appareil de Golgi depuis que LRRK2 y est localisée. Cette localisation ainsi que sa présence dans les vésicules synaptiques et la membrane de la cellule a suggéré une interaction de cette protéine avec les membranes, confirmée par son interaction avec les lipides au niveau de radeaux lipidiques<sup>129</sup>.

Cette protéine joue également un rôle dans la transduction du signal notamment par son domaine kinase qui lui confère un rôle de protéine enzymatique catalysant des réactions de phosphorylation et le domaine ROC associé à une activité GTPase qui permet l'hydrolyse du GTP/GDP<sup>130</sup>. Des études *in vitro* ont démontré que les mutations de LRRK2 notamment les mutations

G2019S et R1441C augmentent l'activité kinase de la protéine, suggérant ainsi un gain de fonction. Cette activité augmentée a été montrée comme étant à l'origine d'une modification de la morphologie neuronale et d'une réduction de la taille des neurites à la fois dans des cultures primaires et dans le système nerveux central de rat *in utero*. Cet effet est inversé par la réintroduction de la forme sauvage montrant un rôle de LRRK2 dans la morphologie des neurites<sup>131</sup>.

LRRK2 est aussi associée à la fonction de dégradation des protéines de la cellule, la mutation G2019S entraînant un raccourcissement des neurites par induction de l'autophagie<sup>132,133</sup>. Un modèle murin invalidé pour le gène *LRRK2* révèle une diminution des mécanismes d'autophagie et des lysosomes contribuant à l'accumulation de l' $\alpha$ -synucléine<sup>134</sup>, également observée chez des souris transgéniques exprimant la forme G2019S de *LRRK2*<sup>127</sup>. Par son activité kinase, elle est capable de phosphoryler des protéines telles que les protéines de la voie MAPK/ERK<sup>135</sup>, ou l'endophiline A1, une protéine impliquée dans la formation des vésicules d'endocytose et interagissant avec des protéines du cytosquelette. Par son domaine ROC, LRRK2 est aussi capable de se fixer à des protéines du cytosquelette telles que la tubuline<sup>136</sup> suggérant son implication dans l'organisation et la polarité cellulaire.

L'expression de la forme sauvage de LRRK2 dans des cellules exposées à un stress oxydant par exposition à de l' $H_2O_2$  induit une augmentation de la survie des cellules par rapport à des cellules exprimant la forme mutée Y1699C ou G2019S. Ceci montre que la forme sauvage de la protéine a un rôle de protection contre le stress oxydant<sup>137</sup>. Des études chez les nématodes montrent que LRRK2 sauvage permet de lutter contre le stress induit par 6-OHDA et empêche la mort des cellules dopaminergiques, à l'inverse de la forme mutée G2019S<sup>138</sup>. L'implication de LRRK2 dans la réponse à ce type de stress passe également par sa capacité à phosphoryler la protéine 4E-BP, un régulateur négatif de la traduction dont l'hyperphosphorylation atténue la résistance des neurones au stress oxydant et leur survie chez la drosophile<sup>139</sup>.

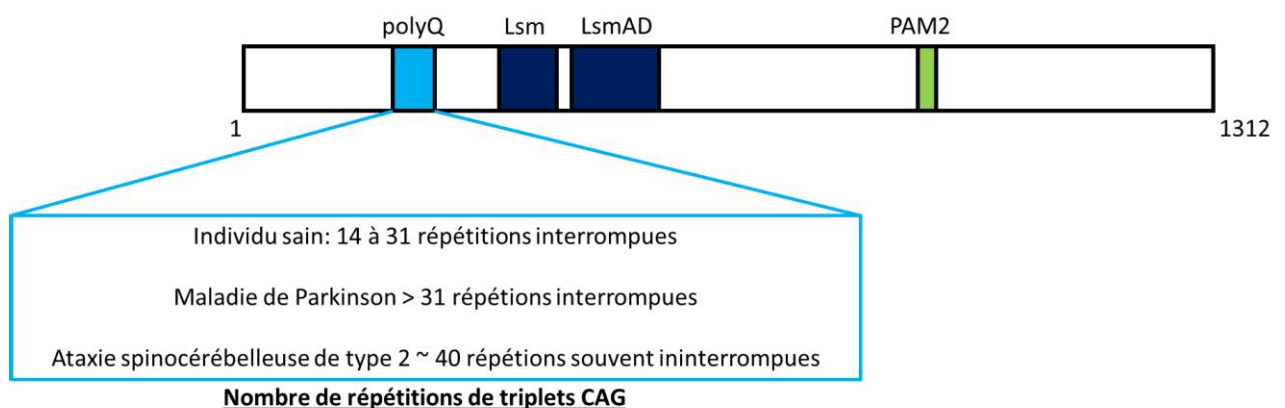
Enfin, le rôle de LRRK2 dans l'inflammation a récemment été suggéré depuis qu'il a été observé que l'inhibition de l'expression de *LRRK2* s'accompagne d'une faible production de molécule proinflammatoire<sup>140</sup> et que cette protéine régule l'activation microgliale après une exposition à la protéine HIV-1 Tat, une protéine inductrice de la réponse inflammatoire<sup>141</sup>. Des polymorphismes de ce gène ont également été associés à des maladies inflammatoires telle que la maladie de Crohn<sup>142</sup>

LRRK2 semble donc être impliquée dans la transduction du signal en particulier via son activité kinase qui permet de moduler différents processus dans la cellule telle que la croissance, la

neurogenèse et le trafic vésiculaire. Elle est également impliquée dans la réponse au stress mitochondrial et très récemment dans la réponse inflammatoire.

### 1.1.4.2.3. Le gène *ATXN2*

Les mutations du gène *ATXN2*, plus souvent connues pour être à l'origine de l'ataxie spinocérébelleuse de type 2 (SCA2), ont aussi été retrouvées chez des patients atteints de MP. Localisé sur le chromosome 12, ce gène s'étend sur une région chromosomique d'environ 130kb. Il comporte 25 exons et se caractérise par la présence d'une portion de séquence dans l'exon 1 contenant des répétitions de triplets de nucléotides CAG. La présence de ces triplets au sein d'un exon se distingue d'autres maladies à triplets comme la dystrophie myotonique par exemple, où les répétitions sont localisées dans des régions non codantes. Ce codon CAG code pour l'acide aminé glutamine dont l'augmentation du nombre conduit lors de la synthèse protéique à une succession de polyglutamine (polyQ). L'allèle physiologique contient 14 à 31 répétitions CAG<sup>143</sup>. Chez les sujets sains, ces répétitions CAG sont interrompues par 1 à 3 triplets CAA (codant également pour un acide aminé glutamine) dont l'organisation peut varier selon les individus (Figure 13).



**Figure 12: Domaines protéiques et répétitions de polyglutamine dans la protéine ataxine-2**

Dans la population générale, plus de 90% portent 22-23 répétitions selon une organisation de type (CAG)<sub>8</sub> CAA (CAG)<sub>4</sub> CAA (CAG)<sub>8</sub>. Chez les sujets malades atteints de MP, cette mutation hétérozygote contient un nombre de répétitions supérieur à 31 avec au moins une interruption CAA. Chez les sujets atteints de SCA2, l'allèle pathologique contient le plus souvent plus de 40 répétitions ininterrompues. Le nombre de répétitions CAG est inversement corrélé à l'âge d'apparition de la maladie et à sa sévérité. Ainsi, la SCA2 qui se déclare classiquement au cours de la troisième décennie peut également apparaître dès l'enfance. Certaines formes juvéniles sont associées à un nombre de répétitions compris entre 130 et 200<sup>144</sup>. Bien que l'augmentation du nombre de triplets CAG dans

*ATXN2* est initialement associée à la SCA2, une augmentation modérée (27 à 33 répétitions) a été associée à un risque accru de développer d'autres maladies neurodégénératives telles que la sclérose latérale amyotrophique (SLA<sup>145</sup>), la dégénérescence lobaire fronto-temporale (DLFT) et la paralysie supra-nucléaire progressive (PSP). Le rôle des triplets CAG interrompus de l'*ATXN2* dans la MP, n'est pas complètement élucidé. Alors que ces triplets ont été identifiés comme la cause de certaines formes familiales de MP, des données contradictoires d'une méta-analyse sur une population caucasienne suggèrent que les sujets ayant un nombre de répétitions supérieur à 32 n'ont pas un risque augmenté de développer une MP<sup>146</sup>. A l'inverse, le seuil de significativité de la longueur des triplets identifiés dans une population japonaise, suggère qu'ils pourraient influencer les résultats de l'association avec la MP ; ce seuil correspondant à un nombre de répétitions égal à 25. Néanmoins l'utilisation de ce seuil dans la méta-analyse n'a pas entraîné de changement drastique de significativité. Aux vues de ces données, il demeure indispensable de mieux comprendre le rôle du gène *ATXN2* dans la MP, dans les formes familiales tout d'abord mais aussi dans les formes sporadiques, ce d'autant que la MP est une maladie dont le caractère multifactoriel est en grande partie relié à des interactions gènes-gènes et gènes-environnements. Une absence d'association pourrait aussi s'expliquer par l'absence de données d'exposition environnementale dans la majorité des études.

Le gène *ATXN2* code un transcrit majeur de 4712pb contenant 25 exons et 2 autres variants d'épissage générés par l'exclusion de l'exon 10 ou de l'exon 21. La découverte, dans l'exon 18, du site de fixation UGCAUG du facteur d'épissage RBFOX2 (aussi connu sous le nom de RBM9 ou HNRBP2) suggère l'existence d'un autre transcrit dont l'épissage serait dépendant de RBFOX2<sup>147</sup>. L'expression de l'*ATXN2* semble être ubiquitaire. Par Northern Blot, Huynh *et al.* ont montré que ce gène était exprimé dans différents tissus cérébraux et non cérébraux<sup>148</sup>. Dans le système nerveux, *ATXN2* est exprimé dans l'amygdale, le noyau caudate, le callosum corpus, l'hippocampus, l'hippocampe, la substance noire, le noyau sous thalamique, le thalamus, le cervelet, la médulla, le cortex cérébelleux, la corde spinale, le putamen, le pôle occipital, le lobe frontal et le lobe temporal. Au niveau du système nerveux central et en particulier dans le cerveau et la corde spinale, le transcrit majeur est prédominant. A l'inverse, le transcrit dépourvu de l'exon 10 est plus abondant dans le cervelet. Le variant d'épissage dépourvu de l'exon 21 est quant à lui exprimé dans le cerveau, la corde spinale, le cervelet, le cœur et le placenta. L'expression de l'*ATXN2* est sous le contrôle de facteurs de transcription tel que Zinc Finger And BRCA1-Interacting Protein With a KRAB Domain 1 (ZBRK1) et v-ets avian erythroblastosis virus E26 oncogene homolog 1 (ETS1)<sup>149</sup>. La protéine ataxine-2 interagit également avec ZBRK1 agissant alors comme un co-activateur transcriptionnel et suggérant un mécanisme d'autorégulation.

La protéine ataxine-2 est une protéine de 140kDa composée de 1312 acides aminés. Elle contient plusieurs domaines protéiques informatifs concernant son rôle physiologique. L'extrémité N-terminale (N-ter) de la protéine contient les domaines polyglutamines (polyQ) et acide allant respectivement des acides aminés 166 à 187 et 280 à 481. L'extrémité C-terminale (C-ter) contient quant à elle un domaine PAM2 (PABP-interacting motif) allant des acides aminés 908 et 925. Outre son domaine acide qui contient les domaines Like Sm Associated Domain (LsmAD) et Like Sm 1 et 2 (Lsm1 et Sm2), la protéine ataxine-2 est hautement basique<sup>150</sup>. Ces domaines Lsm sont des motifs présents sur les protéines liées à l'épissage et sont partagés par les small nuclear ribonucleic proteins (snRNPs)<sup>151</sup>. L'ataxine-2 est essentiellement cytoplasmique et fortement exprimée dans les cellules de Purkinje. Cependant, elle a également été retrouvée dans le réticulum endoplasmique rugueux (RER), le réseau trans-golgien, les granules de stress, les P-bodies et à plus faibles proportions dans le noyau<sup>152</sup>. La localisation de l'ataxine-2 est modulée par des modifications post-traductionnelles telles que la troncation, la phosphorylation et l'ubiquitination. En effet, la phosphorylation de l'ataxine-2 a été retrouvée dans des lignées cellulaires de SY5Y et est impliquée dans le clivage protéolytique de la protéine et/ou son agrégation. Sous sa forme ubiquitinylée, l'ataxine-2 se retrouve dans les agrégats protéiques intranucléaires.

Bien que la fonction de l'ataxine-2 ne soit pas clairement identifiée, sa haute conservation à travers les espèces a rendu possible l'étude de sa fonction dans différents organismes modèles. Plusieurs études ont ainsi révélé que l'ataxine-2 jouait un rôle majoritaire dans le métabolisme des ARNs, en particulier l'épissage, la traduction et la dégradation des ARNs. Elle semble également intervenir dans des processus moins décrits tels que l'endocytose et la dégradation protéique, le métabolisme énergétique ou encore la signalisation calcique.

L'implication de l'ataxine-2 dans l'épissage réside dans la présence de domaines Lsm1 et 2 dans sa région N-terminale, qui sont des domaines caractéristiques des protéines Sm qui constituent les small nuclear Ribonucleoproteins (snRNPs). L'interaction de l'ataxine-2 avec des facteurs liés à l'épissage est également en faveur de son implication dans cette étape de maturation des ARNs. L'identification chez la levure, de l'interaction physique entre l'extrémité C-terminale de l'ataxine-2 et la protéine Ataxin-2 Binding Protein 1 (A2BP1 ou RBFOX1)<sup>153</sup> corrobore cette hypothèse aux vues du rôle de A2BP1 dans l'épissage au sein des cellules neuronales. D'autres facteurs d'épissage interagissent avec l'ataxine-2 tels que RNA-Binding Motif Protein 9 (RBM9)<sup>154</sup>, RNA-Binding Protein With Multiple Splicing (RBPMS), TAR DNA-binding protein 43 (TDP-43) et Fused in sarcoma (FUS)<sup>144</sup>. Il a été démontré que l'expansion de polyQ dans l'ataxine-2 contribuait à une mauvaise localisation des protéines TDP-43 et FUS. De plus, ces expansions de polyQ aggravent la toxicité induite par les

mutations des gènes *TARDBP* et *FUS* contribuant à la physiopathologie de la SLA et la DLFT, deux affections associées à des défauts d'épissage.

Outre son rôle dans l'épissage, l'interaction de l'ataxine-2 avec la protéine PolyA-Binding Protein Cytoplasmic 1 (PABPC1) fait le lien entre l'ataxine-2 et diverses autres fonctions telles que la polyadénylation des ARNm, leur dégradation par le Nonsense-mediated mRNA decay (NMD) ainsi que la traduction. Cette interaction fait intervenir le domaine PAM2 de l'ataxine-2 et le domaine MLE (Mademoiselle) de PABPC1<sup>155</sup>. Les protéines PABPs possèdent des motifs de reconnaissance des ARNs (RRM) qui leur confèrent une haute affinité pour la queue polyA. De cette manière, elles protègent l'ARNm nouvellement synthétisé des attaques par les enzymes de dégradation et contrôlent ainsi leur stabilité. Chez la levure, la délétion des orthologues de PABPC1 et de l'ataxine-2 induit respectivement une augmentation et une diminution de la taille de la queue polyA<sup>144</sup>. L'impact de ces délétions sur la longueur de la queue polyA illustre le rôle de ces protéines dans la stabilité des ARNs, rôle partagé par l'ataxine-2 qui joue sur cette stabilité en se fixant sur l'extrémité 3' non traduite des ARNm<sup>156</sup>. Cette liaison promeut ainsi leur stabilité et module leur niveau de traduction<sup>156,157</sup>

Dans la cellule, certains des ARNs destinés à être dégradés sont localisés dans des sous-compartiments cytoplasmiques appelés les Processing-bodies (P-bodies). Ces P-bodies constituent un lieu de stockage des ARNs qui sont soit en latence soit destinés à être dégradés. Ils contiennent un large lot de protéines associées au métabolisme des ARNs telles que des protéines d'élimination de la coiffe (DCP1 (Decapping protein 1) et DCP2 (Decapping protein 2)), des activateurs des protéines DCPs, des exonucléases 5'-3', des protéines du NMD, des éléments de la voie des microARNs (miARNs) et également des facteurs de la traduction. Il a été montré que l'ataxine-2 était localisée dans ces compartiments. Là, elle interagit avec la protéine DEAD/H-box RNA helicase 6 (DDX6) qui contribue au déroulement des ARNs mais aussi à leur traduction<sup>158</sup>. Cette interaction se fait via les domaines Lsm/LsmAD de l'ataxine-2 et appuie l'implication de l'ataxine-2 dans la stabilité et la dégradation des ARNs. La surexpression, dans différentes lignées cellulaires humaines, de l'ataxine-2 sauvage (Q22) ou mutée (Q79) entraîne une diminution du nombre de P-bodies et perturbent leur assemblage soulignant l'importance de l'ataxine-2 dans les fonctions relatives à ces sous-compartiments cellulaires. La mise en évidence de l'interaction de l'ataxine-2 avec les ribosomes dans les ovaires de drosophile et dans des lignées cellulaires HEK293T ainsi que son interaction avec la protéine PABPC1 est en faveur de son rôle dans la traduction<sup>157,159,160</sup>. PABPC1 interagit via ses domaines RRM avec la protéine eukaryotic Initiation Factor 4G (eIF4G). eIF4G se lie à la coiffe 5' des ARNm et par son interaction avec PABPC1 liée à la queue polyA, permet une circularisation de l'ARNm favorisant ainsi la traduction. La découverte chez la drosophile d'une interaction directe entre l'ataxine-2 et les

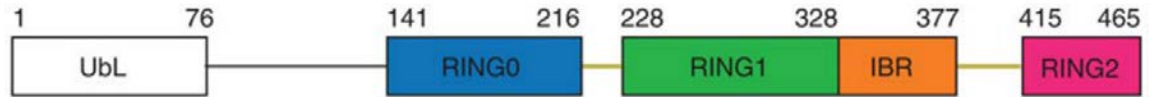
protéines GW182 et Argonaute 1 suggère que l'ataxine-2 pourrait réprimer la traduction via les miARNs, GW182 et Argonaute 1 étant impliquées dans la biosynthèse de miARNs<sup>161</sup>.

Chez la levure, l'ataxine-2 interagit avec l'endophiline A1 et A3, deux protéines interagissant avec les phospholipides membranaires et impliquées dans l'endocytose<sup>162</sup>. La génération d'un modèle murin invalidé pour le gène l'*ATXN2* entraîne une importante prise de poids qui s'accompagne d'une augmentation du niveau d'insuline dans le sérum, témoin d'une insulino-résistance caractéristique des syndromes métaboliques<sup>163</sup>. Celle-ci est suivie d'une augmentation des lipides hépatiques et sanguins également retrouvée au niveau du cervelet. L'ensemble de ces modifications induites par la déplétion de l'*ATXN2* lie ainsi l'ataxine-2 au métabolisme et en particulier au métabolisme lipidique et glucidique. La localisation de l'ataxine-2 dans le RE, principal lieu de stockage du calcium intracellulaire, suggère l'intervention de cette protéine dans le métabolisme calcique associé à la pathogénèse de la SCA2. Jing Liu et ses collaborateurs ont démontré chez la souris que la protéine ataxine-2 (Q58) interagissait avec l'extrémité C-terminale de la protéine Inositol 1,4,5-trisphosphate Receptor type 1 (InsP3R1), un canal calcique fortement exprimé dans le RE des cellules de Purkinje<sup>163,164</sup>. Cette interaction conduit à une sensibilité accrue du canal calcique à son activateur l'inositol 1,4,5-trisphosphate (InsP3) conduisant à une augmentation du calcium intracellulaire pouvant être délétère.

#### 1.1.4.2.4. Le gène *PRKN*

Le gène *PRKN* situé au locus 6q26 est le premier gène identifié comme étant responsable de formes génétiques à transmission autosomique récessive à début précoce. Ce gène s'étend sur une région de 1,53Mb et contient 12 exons<sup>165</sup> le long desquels ont été identifiés plus de 120 mutations, délétions et multiplications<sup>166</sup>. Les sujets porteurs de mutations de ce gène déclarent la maladie vers 40 ans et parfois même durant l'adolescence ; ces formes de la MP rendant compte d'environ 50% des cas de MP à début précoce.

Ce gène code la parkin, une protéine de 465 acides aminés (figure 13) qui contient un domaine ubiquitine-like (Ubl) en N-terminale et de 4 domaines RINKG-like (RING0, RING1, RING2 et IBR)<sup>167</sup>. Elle se caractérise par une activité E3 ubiquitine ligase, une enzyme impliquée dans la dégradation des protéines dans la cellule. Ce rôle dans la dégradation des protéines requiert l'interaction de la parkin avec son partenaire, la protéine E2 qui catalyse l'ajout d'ubiquitine sur les résidus lysines des protéines, ce qui constitue un signal pour leur dégradation. L'interaction de la parkin avec l'enzyme E2 est annulée par la mutation T240M.



**Figure 13: Domaines protéiques de la parkin (d'après Riley & Johnston, 2012)**

Le rôle majeur de la parkin se situe dans le contrôle de l'homéostasie mitochondriale. Elle agit de concert avec la protéine PINK1 qui permet le recrutement de la parkin au niveau des mitochondries endommagées. Ce recrutement permet l'ubiquitination de la membrane externe de la mitochondrie, constituant un signal pour sa dégradation par mitophagie (autophagie de la mitochondrie). Plusieurs mutations pathologiques de la parkin empêchent la dégradation des mitochondries endommagées entraînant une augmentation du stress oxydant<sup>168</sup>. La parkin peut également ubiquitiner des protéines de la membrane externe telles que les mitofusines, la protéine OPA1 ou encore la protéine Miro qui jouent un rôle dans l'équilibre fusion/fission, la motilité et permettent d'isoler les mitochondries à éliminer.

La parkin agit également en dehors de la mitochondrie. Sa localisation au niveau des vésicules riches en protéines oxydées suggère un rôle de la parkin dans le trafic vésiculaire<sup>169</sup>. Elle intervient également dans la régulation d'expression génique en interagissant avec le facteur de transcription mitochondrial TFAM. Enfin, elle prévient la mort cellulaire en entraînant la dégradation de la protéine PARIS (parkin interacting substrat), un répresseur de l'expression du facteur de transcription PGC1- $\alpha$  qui contrôle l'expression de gènes de la chaîne respiratoire<sup>170</sup>. La délétion de la parkin dans des modèles animaux entraîne la mort de cellules neuronales dopaminergiques. Ce rôle dans la mort cellulaire est étayé par son interaction avec la protéine Bcl et le fait que des délétions de ce gène soient observées dans 30% des cancers humains, contribuant à montrer que ce gène est un suppresseur de tumeurs.

#### 1.1.4.2.5. Le gène *PINK1*

Le lien entre *PINK1* (PTEN-induced putative kinase 1) est la MP a été évoqué en 2001 par la mise en évidence d'une mutation à transmission autosomique récessif au sein d'une famille italienne consanguine<sup>171</sup>. Plus de 50 mutations faux-sens de *PINK1* ont été identifiées, qui après les mutations de la parkin, constituent la seconde cause de MP récessive et représentent 1-8% des formes monogéniques conduisant à une forme de MP atypique avec une apparition précoce des premiers symptômes et une progression plus lente de la maladie.



Ce gène constitué de 8 exons code une protéine cytoplasmique de 63 kDa composée de 581 acides aminés qui définissent 3 domaines principaux : une séquence N-terminale allant du premier acide aminé au 77<sup>ème</sup> qui permet l'adressage de la protéine à la mitochondrie, un domaine transmembranaire allant de l'acide aminé 94 à 110 et un domaine C-terminal allant de l'acide aminé 156 à 511 qui porte une activité sérine/thréonine kinase. Cette protéine est largement exprimée dans le cerveau notamment au niveau du cortex temporal, de l'amygdale, de la substance noire, dans le cervelet et dans le noyau denté<sup>172</sup>.

La plupart des mutations de PINK1 se situent au niveau du domaine kinase et entraîne une inhibition de son activité de phosphorylation<sup>173</sup>. La perte de fonction de PINK1 est associée à une accumulation de mitochondrie défectueuse pouvant être à l'origine de la surproduction de ROS et du stress oxydant<sup>174</sup>. Cette production accrue est associée à une inhibition de l'activité du complexe 1 de la chaîne respiratoire mitochondriale<sup>175</sup> suggérant que cette perte de fonction joue un rôle crucial dans la pathogenèse de la MP<sup>176</sup>. Du fait de la présence du signal d'adressage à la mitochondrie dans son domaine N-terminal, PINK1 est consitutivement adressée à la mitochondrie. Elle agit comme un senseur des dommages de la mitochondrie. En situation physiologique, PINK1 est transloquée de la membrane externe vers la membrane interne de la mitochondrie à travers les complexes de translocation mitochondriaux TOM (Translocon of the Outer Membrane) et TIM (Translocon of the Inner Membrane). Une fois à la membrane interne, elle est clivée par les protéines MPP et PARL entre les acides aminés 103 et 104. Ce clivage libère un fragment de 52kDa qui va être relargué dans le cytoplasme et interagir avec la parkin bloquant ainsi sa translocation à la mitochondrie<sup>177</sup>. Lorsque les mitochondries sont endommagées, le clivage de PINK1 est inhibé annulant le blocage de la parkin dans le cytoplasme qui peut alors être transloquée à la mitochondrie et induire la dégradation des mitochondries endommagées<sup>178</sup>. Il a été montré *in vitro*, que PINK1 était capable de phosphoryler la parkin au niveau de la serine 65. Cette phosphorylation est détectée lorsque que la mitochondrie est dépolarisée et semble être un mécanisme d'induction de la mitophagie.

#### 1.1.4.2.6. Le gène DJ-1

Le lien entre DJ-1 et la MP a été identifié par une analyse génomique au sein d'une famille hollandaise consanguine et a été confirmé par la suite au sein d'un famille italienne<sup>179</sup>. Plusieurs mutations (délétions d'exons, mutations faux-sens) ont été identifiées au sein de ce gène comptant pour moins de 1% des formes monogéniques précoces.

Ce gène code la protéine DJ-1 de 189 acides aminés et dont l'expression est ubiquitaire. Elle est retrouvée aussi bien dans les neurones que dans les astrocytes. Dans la cellule, elle est localisée

au niveau du cytoplasme ainsi que dans le noyau et dans la mitochondrie. Cette protéine multifonctionnelle joue un rôle majeur dans la gestion du stress oxydant. En effet, la perte de fonction de DJ-1 a été montrée comme induisant une augmentation de la production de ROS ainsi qu'une réduction du potentiel de membrane mitochondriale<sup>180</sup>. De plus, la sensibilité au stress induit par le MPTP est plus important dans les astrocytes exprimant une protéine DJ-1 déficiente<sup>181</sup>. Dans les lignées cellulaires dopaminergiques de rat N27 et dans les neurones primaires dopaminergiques, la surexpression de la forme DJ-1 sauvage protège les cellules de la mort induite par l'H<sub>2</sub>O<sub>2</sub> et la 6-OHDA à l'inverse de la forme mutée DJ-1 L166P (la plus étudiée)<sup>182,183,184</sup>. Elle agit donc comme agent antioxydant dont le mode d'action semble reposer sur sa liaison aux ions cuivre au niveau de la cystéine 106 permettant le recrutement de l'enzyme anti-oxydante SOD1<sup>185</sup>.

L'équilibre entre fusion et fission mitochondriale est perdu en présence des formes mutées G309D et C106A de DJ-1 impliquant la protéine dans la régulation de l'homéostasie mitochondriale<sup>186</sup>. La perte de cet équilibre induit par la perte de fonction de PINK1 est restaurée par l'expression de la forme sauvage, mais aussi par les protéines DJ-1 et parkin, montrant que ces protéines ont des fonctions communes. La perte de fonction de DJ-1 est aussi associée à une réduction du potentiel de membrane et une accumulation de marqueurs autophagiques<sup>187</sup> suggérant un rôle dans les mécanismes de dégradation de la cellule.

D'autres fonctions de DJ-1 sont peu documentées comme son rôle dans la régulation de l'expression génique. Dans les lignées cellulaires dopaminergiques de neuroblastome SK-N-SH, la surexpression de la protéine sauvage DJ-1 est capable de séquestrer le répresseur de la transcription de la TH, le facteur PSF. Cette séquestration est inhibée par l'invalidation de DJ-1<sup>188</sup>. DJ-1 est aussi capable de s'associer avec les radeaux lipidiques via la palmitoylation de ses cystéines C46, C53 et C106 et sa région C-terminale<sup>189</sup>.

La protéine DJ-1 joue un rôle dans la mort cellulaire en se liant à la protéine mitochondriale Bcl-XL régulatrice de l'apoptose. Cette liaison étant renforcée par la mutation L166P entraînant une dissociation de la protéine apoptotique BAX initialement fixée à Bcl-XL conduisant à la mort<sup>190</sup>.

#### 1.1.4.2.7. Le gène *VPS35*

La liaison entre VPS35 (Vacuol protein sorting 35) et la MP a récemment été identifiée. Par une technique de séquençage des exons, deux équipes ont mis en évidence au sein de ce gène, une mutation à transmission autosomique dominante qui conduit à un phénotype semblable à celui des formes sporadiques. Cette mutation correspond à un changement de la guanine en position 1858 par

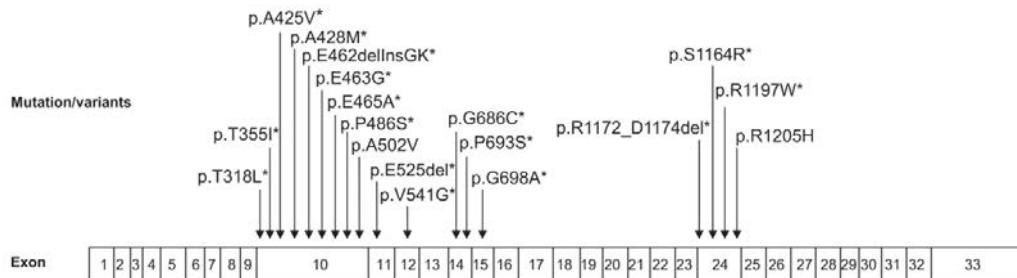
un nucléotide adénine qui conduit au niveau de la séquence protéique à un changement de l'acide aspartique en une asparagine en position 620 (p.Asp620Asn)<sup>191,192</sup>. D'autres variants ont été identifiés mais le rôle pathogène n'a pas encore été démontré.

Situé au locus 16q11.2, ce gène code pour une protéine faisant partie du complexe rétromère. Chez l'homme ce complexe est composé des protéines SNX1, SNX2, Vps26, Vps29 et Vps35<sup>193</sup>. Il intervient dans le trafic des vésicules permettant l'adressage de protéines du réseau trans-golgien vers le lysosome. L'inhibition de Vps35 est associée à une diminution du pool de protéines cargots dans le réseau trans-golgien et à leur accumulation dans les endosomes confirmant son implication dans le transport des protéines<sup>194</sup>. Son rôle dans le trafic vésiculaire passe également à travers son interaction avec les protéines SNARE<sup>191</sup>. Chez la drosophile, la perte de fonction de l'homologue de Vps35 entraîne une mauvaise localisation d'un grand nombre de récepteurs et à une augmentation de la prolifération des cellules sanguines suggérant également un rôle dans la prolifération cellulaire. Cette inhibition cause également une activation de la formation des terminaisons synaptiques et la protéine inhibe la polymérisation de l'actine<sup>195</sup>. La mutation D620N entraîne une perte d'interaction de la protéine Vps35 avec les protéines FAM21, un constituant du complexe WASH situé à la surface des endosomes et impliqué dans la polymérisation de l'actine. De plus, l'inhibition dans des cellules HEK293T entraîne des perturbations de la communication entre le peroxysome et la mitochondrie, deux organites participant à la respiration cellulaire, pouvant affecter le stress oxydant<sup>196</sup>. La surexpression de Vps35 dans des lignées cellulaires dopaminergiques de rat N27 les protège de la toxicité induite par le MPTP faisant aussi le lien entre Vps35 et la mitochondrie<sup>197</sup>. Enfin, Vps35 semble jouer un rôle dans les mécanismes de dégradation protéique par son interaction avec la protéine parkin et par l'augmentation de la protéine ubiquitine ligase E3 induite par la mutation D620N<sup>198</sup>. Ceci étant confirmé par l'inactivation de Vps35 qui induit une accumulation de l' $\alpha$ -synucléine dans le cerveau de drosophile déplétées<sup>199</sup>.

#### 1.1.4.2.8. Le gène *EIF4G1*

En 2011, Chartier-Harlin et ses collaborateurs ont mis en évidence une mutation faux-sens chez des patients parkinsoniens venant d'une famille du nord de la France<sup>200</sup>. Cette mutation correspond à un changement au niveau du nucléotide 3614 d'une guanine en adénine (G3614A) induisant le changement de l'arginine en position 1205 en une histidine au niveau de la protéine (p.R1205H) (Figure 14). Depuis cette découverte, le séquençage d'exome complet n'a pas permis d'identifier d'autres mutations.

La mutation p.R1205H est localisée dans le 24<sup>ème</sup> exon, ségrège avec la maladie de manière autosomique dominante et induit chez les patients une MP à début asymétrique et marquée par la présence de corps de Lewy<sup>200</sup>.



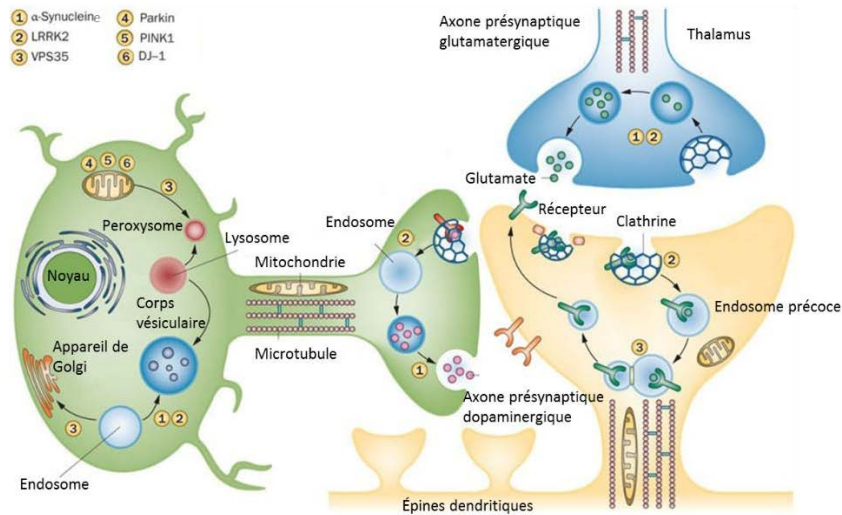
**Figure 14: Représentation de la protéine EIF4G1 et de ses mutations (d'après Deng, Wu & Jankovic, 2015)**

Bien qu'initialement décrite comme une mutation présente à la fois chez les parkinsoniens familiaux et sporadiques, cette mutation a également été identifiée chez les témoins avec une fréquence pouvant être supérieure à celle observée chez les malades<sup>201,202</sup>.

Ce gène code une protéine de 1599 acides aminés qui joue un rôle dans la traduction notamment par son interaction avec les protéines PABP, eIF4E, eIF4A, eIF3 et Mnk1 qui contribuent à l'initiation de la synthèse protéique<sup>203,201</sup>. EIF4G1 intervient également dans la régulation de l'expression par le biais des miARNs, en interagissant avec la protéine Ago2, un constituant du complexe RISC qui permet la reconnaissance des ARNs cibles par les miARNs<sup>204</sup>. Très récemment, Dhungel *et al.* ont montré que EIF4G1 était capable d'interagir avec d'autres protéines dont les mutations de gènes sont liées à la MP, comme la protéine Vps35 et l' $\alpha$ -synucléine<sup>205</sup> suggérant un rôle d'EIF4G1 dans les mécanismes de dégradation protéique<sup>206</sup>. Cependant l'implication d'EIF4G1 dans la MP est actuellement remis en cause depuis que la mutation p.R1205H a été retrouvée plus souvent chez les sujets sains que chez les malades<sup>202,207</sup>. Aux vues de ces nouvelles données, notre équipe en collaboration avec M. Farrer envisage une réévaluation clinique des patients chez qui nous avons identifié la mutation p.R1205H ainsi qu'un séquençage de l'ensemble du génome afin d'identifier si d'autres mutations sont présentes chez ces sujets pouvant expliquer leur phénotype parkinsonien.

En résumé, l'identification des facteurs environnementaux et génétiques liés à la MP a grandement contribué à l'amélioration des connaissances des mécanismes moléculaires de la maladie puisqu'ils ont permis de mettre en évidence les perturbations cellulaires induites par ces facteurs. Ainsi, il ressort des différentes études que les dysfonctionnements de mitochondrie,

l'induction du stress oxydant, l'inhibition des mécanismes de dégradations protéiques, les perturbations du transport vésiculaire et l'inflammation sont des mécanismes contribuant à la pathogenèse de la MP (figure 15).



**Figure 15 : Mécanismes moléculaires physiopathologiques de la maladie de Parkinson (d'après Trinh & Farrer, 2013)**

*Ce schéma illustre les mécanismes moléculaires de la MP mis en évidence par l'étude de certains des gènes responsables des formes génétiques de la maladie : l'analyse de fonctions et des perturbations engendrées par les mutations de SNCA ayant permis de révéler son implication notamment dans le trafic vésiculaire (endocytose/exocytose) et dans la transmission dopaminergique ; LRRK2 qui elle est liée à la transduction du signal via sa fonction kinase mais qui a également la capacité d'interagir avec les membranes permettant ainsi le flux de vésicules ; de même pour Vps35 qui par sa fonction dans le complexe retromer assure le transport des endosomes à l'appareil de Golgi et enfin les protéines PINK1, DJ-1 et parkin qui agissent de concert au niveau de la mitochondrie assurant sa dégradation lorsqu'elle est endommagée sachant que la parkin a aussi un rôle dans la dégradation par le lysosome.*

La diversité génétique de la MP montre que de nombreux mécanismes peuvent être à l'origine de la maladie. Malgré l'identification de plusieurs d'entre eux, les traitements actuellement disponibles restent incapables d'agir sur la progression de la maladie. Ceci suggère que la connaissance des mécanismes pathogènes est encore incomplète et que d'autres restent encore à découvrir.

#### 1.1.4.3. L'apport des études d'association

L'essor des outils d'analyse à haut débit permet le séquençage rapide et simultané de nombreux échantillons. Basées sur ces technologies, les études d'association génétique et les études pangénomiques de type GWAS (Genome-wide association study) contribuent à l'identification des polymorphismes nucléotidiques associés à une maladie. Suite au développement de ces

techniques, plusieurs études d'association ont été menées dans le cadre de la MP par l'utilisation de puces à SNP (Single Nucleotide Polymorphism) révélant les polymorphismes fréquemment retrouvés chez les malades. Les premières datent du début des années 2010 et ont permis d'identifier de nombreux facteurs de risque au sein des gènes connus pour leur implication dans la MP tels que *SNCA* et *LRRK2* mais aussi dans le gène *MAPT* impliqué dans la polymérisation des microtubules et faisant le lien entre le cytosquelette et la MP<sup>208</sup>. Le cluster de SNP associé à la région 3'UTR de *SNCA* suggère que le variant causal pourrait affecter la maturation post-transcriptionnel de l'ARNm de *SNCA* en influant sur la liaison des miARNs qui régulent son expression ou par un effet sur son épissage alternatif<sup>208</sup>. La plupart des études d'association ont été menées au sein de populations parkinsoniennes présentant une forme sporadique de la maladie. Toutefois, ces polymorphismes sont aussi identifiés chez les sujets présentant des formes familiales qui ont permis de révéler le risque associé aux polymorphismes des gènes *DGKQ* et *GAK*<sup>209,210</sup>. Le locus PARK16 contenant le gène *BST1* a été montré comme renfermant une susceptibilité accrue à la MP dans les populations asiatiques<sup>211</sup> et caucasiennes<sup>212</sup>.

Les études d'association sont informatives mais le nombre d'individus utilisés pour ce type d'études reste un facteur limitant<sup>213</sup>. De fait, l'identification de variants rares est compromise. Le développement des méta-analyses regroupant les données de GWAS effectuées dans le cadre de l'étude d'une même maladie a permis d'obtenir des informations sur un plus grand nombre de sujets et d'augmenter la robustesse des polymorphismes identifiés. Concernant la MP, les polymorphismes les plus robustes c'est-à-dire les plus communément retrouvés sont localisés dans des gènes déjà connus pour avoir un rôle dans la MP tels que *SNCA*, *LRRK2* mais également dans les gènes *MAPT*, *GAK*, *RIT2*<sup>214</sup>, *BST1* et *HLA-DRB5* appuyant l'importance de l'immunité dans la MP<sup>210,215,216,217</sup>. Le gène *GBA* responsable de la maladie de Gaucher, une maladie lysosomale dont certains patients présentent un syndrome parkinsonien, semble à l'heure actuelle être le plus grand facteur de susceptibilité de la MP liant les mécanismes de dégradation cellulaire à la MP, de même que les polymorphismes observés dans le gène *LAMP3* codant pour une protéine de la membrane des lysosomes<sup>218</sup>. Dernièrement, Nalls *et al.* ont utilisé les données de 15 GWAS permettant l'identification de 26 variants génétiques associés à la MP incluant des loci à risque déjà identifiés tels que *GBA-SYT11*, *TMEM175-GAK-DGKQ*, *SNCA*, *MAPT* et 22 autres ayant avec une forte probabilité d'association.

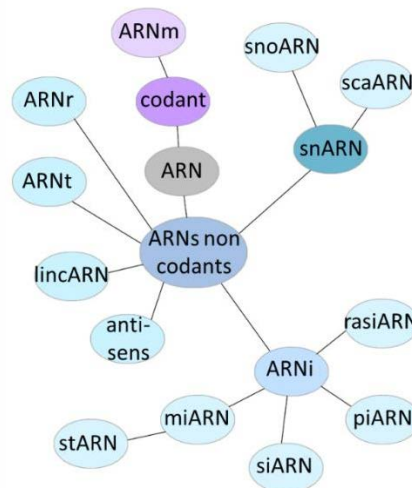
Alors que ces études d'association ont permis de révéler des polymorphismes au sein de gènes connus de la maladie, elles ont conduit à en suspecter d'autres dont le rôle causal n'est pas encore déterminé et reste à démontrer. Alors que l'hérédité de la MP est de 27%, les formes

monogéniques n'expliquent que 10% des formes familiales suggérant que la part des facteurs de susceptibilité est importante dans l'induction de la maladie. Face à la grande diversité des mécanismes moléculaires potentiels de la MP, le recours à des outils d'analyse à grande échelle semble inévitable dans la recherche de ces nouveaux mécanismes.

Plusieurs techniques d'analyse globale ont été développées et reposent sur l'analyse des gènes (génomique), des ARNs (transcriptomique), des protéines (protéomique) ou encore des différents métabolites cellulaires (métabolomique). Nous allons nous intéresser plus particulièrement au transcriptome qui suscite un grand intérêt au sein de notre laboratoire.

#### 1.1.4.4. La pertinence de l'étude du transcriptome dans la maladie de Parkinson

L'information génétique portée par le génome nucléaire se transmet à l'ensemble de la cellule par divers mécanismes permettant son expression. Le premier est la transcription qui va permettre la conversion des gènes contenus dans le génome en une molécule d'ARN appelé transcrite. L'ensemble des transcrits constitue le transcriptome dont la diversité est représentée ci-dessous (figure 16).



**Figure 16 : Différents types d'ARNs codés par le génome humain.**

*Lors de la transcription, les gènes présents dans le génome humain vont être convertis en molécules appelées ARNs dont on distingue plusieurs catégories: les ARNs codants et les ARNs non codants. Les ARNs codants appelés ARN messagers (ARNm) vont servir de matrice lors de la traduction de ces ARNm en protéines. Les ARNs non codants se subdivisent en plusieurs sous-catégories incluant les ARN ribosomiques (ARNr) et ARN de transfert (ARNt) qui constituent la machinerie effectrice de la traduction, les ARNs small nucleolar ARNs (snoARNs) et les small nuclear ARNs (snARNs) impliqués dans la maturation des ARNr, les ARNs small cajal body-specific (scaARNs) impliqués dans la*

*maturation des sous-unités de la machinerie effectrice de l'épissage et enfin les ARNs impliqués dans la régulation de l'expression génique. Cette dernière catégorie regroupe un grand nombre de petits ARNs interférents (ARNi) tels que les micro-ARNs, les siARNs, les piARNs, les rasiARNs et les stARNs qui sont impliqués dans la régulation de l'expression génique, de même que les lincARNs qui sont de taille plus importante*

Le transcriptome reflète l'ensemble des gènes qui s'expriment dans la cellule et permet d'étudier ceux dont l'expression conduit à l'apparition d'un phénotype (caractère donné). L'expression est régulée par des mécanismes épigénétiques tels que 1) la méthylation de l'ADN (mécanisme le mieux caractérisé) au niveau des îlots CpG dans le promoteur ou à l'intérieur d'un gène modulant la fixation des facteurs de transcription, 2) la condensation de la chromatine liée aux modifications post-traductionnelles des protéines histones (méthylation, l'acétylation des lysines, la phosphorylation ou la SUMOylation) qui induisent des changements de conformation et de la structure de la molécule d'ADN génomique à laquelle elles sont associées et enfin 4) les ARNs non codants qui conduisent à la dégradation et/ou à l'inhibition de la traduction des ARNm avec lesquels ils interagissent (c'est le cas des miARNs). Ils jouent également sur l'accessibilité des gènes aux facteurs de transcription en modulant leur fixation, en les séquestrant ou en changeant la structure de la chromatine (c'est le cas des ARNs long non codants).

Ces dernières années, les données de la littérature ont fait mention de nombreuses perturbations de ces mécanismes dans la MP. En effet, l' $\alpha$ -synucléine est capable de modifier le profil de méthylation de la cellule en séquestrant la méthyltransférase DNMT1 dans le cytoplasme ; ceci réduit ainsi l'accès de cette enzyme aux molécules d'ADN présentes dans le noyau et module l'expression de ses cibles<sup>94</sup>. Une hypométhylation de l'intron 1 du gène de l' $\alpha$ -synucléine a été observée dans le sang<sup>90</sup> et le cerveau des patients atteints de la MP où elle est associée à un fort niveau d'expression de l'ARN et de la protéine<sup>219</sup>. L' $\alpha$ -synucléine est aussi capable de se fixer aux protéines histones<sup>220</sup>, ce qui masque leurs sites d'acétylation<sup>221</sup> et a pour conséquence une condensation de la chromatine qui inhibe l'expression de nombreux gènes. *PGC-1 $\alpha$* , un facteur de transcription mitochondrial, est l'un des gènes dont le changement des niveaux de méthylation peut conduire à des modifications assimilables à des perturbations connues dans la MP ; l'hypométhylation au niveau de son promoteur conduit à des dysfonctions de la mitochondrie<sup>222</sup>.

Les ARNs non codants sont également affectés au cours de la maladie ; c'est le cas notamment de miR-7, miR-153 ou encore miR-34b/c capables de fixer l'extrémité 3' non traduite de *SNCA* et de réprimer son expression. La sous-expression de miR-7 et miR-153 dans la substance noire de souris exposées au MPTP est associée à une surexpression de l' $\alpha$ -synucléine<sup>96</sup>. D'autres gènes impliqués dans la MP sont régulés par des micro-ARNs. *LRRK2*, par exemple, est la cible de miR-205,



un micro-ARN qui est sous-exprimé chez les sujets présentant une forme sporadique de la maladie qui en conséquence présente un haut niveau d'expression de LRRK2<sup>124</sup>. De plus, LRRK2 a été montrée comme étant capable d'interagir avec les protéines argonautes, essentielles à la biogenèse des micro-ARNs<sup>223</sup>, suggérant une forte modification du profil d'expression de ces derniers pouvant affecter l'expression d'un grand nombre de cibles. Enfin, certains miARNs comme miR-4697 présentent des polymorphismes associés à la MP suggérant l'importance de la régulation par les miARNs dans la maladie<sup>224</sup>.

Ces données montrent le bien-fondé de l'étude du transcriptome dans la MP puisque ces variations peuvent entraîner des changements d'expression de transcrits au sein de processus cellulaires dont les perturbations sont à l'origine de la MP. L'analyse du transcriptome devrait alors permettre d'identifier ces variations et d'élargir les connaissances des mécanismes à l'origine de la maladie.

#### 1.1.4.5. L'apport du transcriptome de cerveaux *post-mortem*

Plusieurs équipes ont d'ores et déjà eu recours à l'étude du transcriptome pour identifier les processus cellulaires pathologiques de la MP. Ces études reposaient sur l'utilisation de micropuces qui, par appariement des transcrits à des sondes complémentaires, permettent une mesure quantitative des niveaux d'expression de transcrits. La comparaison de ces niveaux dans différents échantillons identifie des variations inhérentes aux différences qu'ils existent entre les échantillons comparés (figure 17).

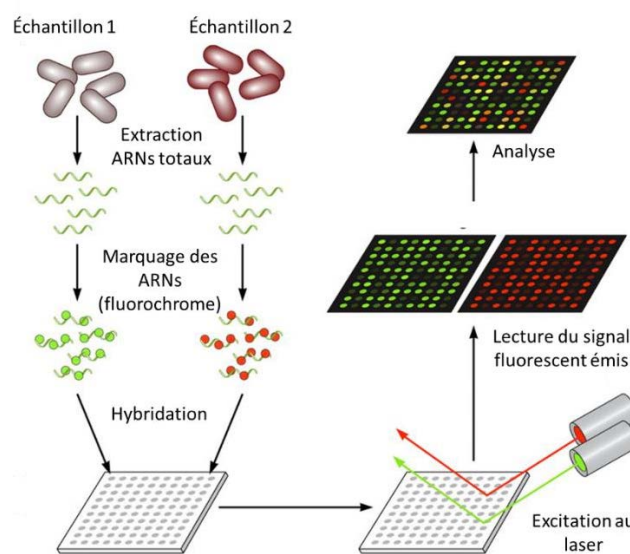


Figure 17 : Principe de l'analyse du transcriptome à l'aide de puce d'expression

Les échantillons 1 et 2 dont on souhaite comparer les niveaux d'expression vont fournir les ARNs nécessaires à l'analyse du transcriptome à l'aide de puces d'expression. Les ARNs de chaque échantillon vont donc être extraits puis marqués de façon différente pour distinguer la provenance de chaque molécule d'ARNs. Ces derniers vont ensuite être déposés simultanément sur un support au niveau duquel se trouvent des sondes complémentaires d'ARNs présents dans les 2 échantillons à analyser. La fixation des ARNs sur les sondes est alors révélée par lecture du marquage préalablement fixé aux ARNs (ici, une fluorescence émise par le fluorochrome après excitation au laser). On a alors deux images de la puce : une correspondant aux signaux émis par les ARNs provenant de l'échantillon 1, l'autre provenant de ceux émis par l'échantillon 2. Il existe une relation de proportionnalité entre le signal émis et la quantité d'ARNs hybridée sur chaque sonde de la puce. Le calcul des ratios de fluorescence entre l'échantillon 1 et l'échantillon 2 permet, pour chaque gène, de déterminer les niveaux d'expression relative

Le cerveau ayant longtemps été considéré comme le principal organe touché, les premières études transcriptomiques ont été basées sur l'étude des ARNs provenant d'échantillons cérébraux de patients. N'étant pas accessible du vivant de ces derniers, ces études reposent sur l'utilisation de tissus *post-mortem*. De nombreux gènes impliqués dans différentes voies métaboliques ont été montrés comme étant dérégulés et pouvant contribuer à la mort neuronale. Grünblatt et ses collaborateurs<sup>225</sup> ont été les premiers à réaliser l'étude du transcriptome à partir de la substance noire *pars compacta* de patients et de témoins appariés. Cette étude a mis en exergue la variation d'expression de gènes impliqués dans différents processus déjà connus pour leur implication dans la MP. En effet, la majorité des gènes sous-exprimés dans cette étude sont liés à des fonctions telles que la transduction du signal, le système de dégradation protéique via le système ubiquitine/protéasome et la signalisation dopaminergique ; les gènes surexprimés étant quant à eux liés à des processus tels que l'adhésion cellulaire, le cytosquelette, le métabolisme protéique, la traduction et le stress cellulaire induit par le biais de l'inflammation. Des résultats complémentaires ont été apportés par Duke *et al.* qui ont montré une dérégulation des gènes de la fonction mitochondriale et du système ubiquitine/protéasome également dans la substance noire de parkinsoniens<sup>226</sup>. La majorité de ces gènes mitochondriaux et de dégradation protéique sont sous-exprimés dans le cerveau de ces malades corrélant avec les pertes de fonctions des gènes des formes monogéniques tels que *PRKN*, *PINK1*, ou encore *DJ-1* et avec l'inhibition de la respiration mitochondriale rapportée à de multiples reprises dans la MP<sup>202</sup>. La substance noire n'étant pas exclusivement composée de neurones dopaminergiques, Simunovic *et al.* ont utilisé la microdissection laser pour isoler les neurones dopaminergiques de la *SNpc* de cas sporadiques afin d'avoir accès au profil d'expression spécifique des cellules majoritairement touchées dans la pathologie<sup>227</sup>. Leurs données transcriptomiques ont montré des résultats similaires à ceux obtenus à partir de la substance noire totale, avec des dérèglements de gènes impliqués dans la fonction mitochondriale, la fonction synaptique, le cytosquelette, les récepteurs aux neurotransmetteurs et la mort cellulaire programmée. De façon intéressante, ils ont également trouvé une diminution de

l'expression des gènes liés aux formes monogéniques tels que *SNCA* et *DJ-1* mettant en avant l'implication de ces gènes dans la pathogenèse des formes sporadiques.

L'inconvénient des études basées sur l'utilisation des tissus *post-mortem* est qu'elles révèlent des dérèglements le plus souvent caractéristiques des stades avancés, la mort neuronale étant importante au moment du décès. La possibilité d'identifier des variations moléculaires précoces à l'origine de la maladie en est alors réduite. De plus, la qualité des ARNs obtenue à partir de ces tissus *post-mortem* est fortement impactée par la longueur du délai *post-mortem* qui corrèle avec une forte dégradation des ARNs et réduit la qualité de l'analyse. La recherche s'est alors tournée vers l'utilisation de tissus dérivés de patients vivants rendant alors possible une étude des premières étapes des processus pathologiques.

#### 1.1.4.6. L'apport du transcriptome de cellules dérivées de patients

Plusieurs tissus accessibles du vivant du patient ont été proposés comme modèles d'étude de la MP. C'est le cas des fibroblastes, des cellules souches pluripotentes induites et des cellules sanguines<sup>228-230</sup> qui à l'heure actuelle constituent la base des études moléculaires dans la recherche de biomarqueurs cliniques de la MP.

L'utilisation de telles cellules a cependant nécessité des validations quant à leur capacité à refléter les modifications moléculaires caractéristiques de la MP. La recherche à partir de l'étude du transcriptome de fibroblastes s'avère prometteuse puisqu'elle a révélé que ces cellules présentaient certains aspects de la maladie notamment par la mise en évidence de troubles mitochondriaux et la variation d'expression de gènes déjà identifiés comme étant induit par le stress oxydant. Ces dérèglements ont été observés dans des fibroblastes de patients porteurs de la mutation G309D de *PINK1* confirmant le lien entre ce gène et la mitochondrie<sup>231</sup>. Leur utilisation est informative puisqu'elle a également permis de montrer l'impact du stress oxydant sur l'induction de l'expression de *SNCA*<sup>231</sup>. Ces fibroblastes dérivés de patients présentent également l'avantage de se différencier en cellules souches (cellules souches pluripotentes induites) qui sous l'effet de facteurs de croissance peuvent être reprogrammées en neurones dopaminergiques présentant le même génotype que le patient dont les fibroblastes sont issus. Ces cellules dopaminergiques dérivées de cellules induites constituent également une source d'information ce d'autant qu'elles permettent d'étudier les changements dans des neurones dopaminergiques qui jusqu'ici n'était possible qu'après le décès du patient. Elles récapitulent certains aspects de la maladie en révélant des variations d'expression de gènes impliqués dans la biosynthèse de la dopamine tels que la sous

expression de la tyrosine hydroxylase, la surexpression de la MAOB ainsi que la surexpression des gènes *SNCA*, *PINK1* et *PRKN*<sup>229</sup>.

Les cellules sanguines ont fait l'objet d'un plus grand nombre d'étude probablement dû au fait que les patients sont plus enclins à subir une prise de sang qu'à la pratique de la biopsie cutanée. De plus, ces cellules présentent un certain nombre de caractéristiques retrouvées au niveau du cerveau de parkinsoniens. En effet, plusieurs données de la littérature montrent une diminution de l'activité des complexes 1 et 3 de la chaîne respiratoire mitochondriale au sein de ces cellules<sup>232-234</sup>. Cette inhibition s'accompagne d'une forte production de ROS, retrouvée également au niveau cérébral<sup>235</sup>. Un rôle de ces cellules dans la pathogénèse de la MP a été proposé depuis qu'elles ont été retrouvées au niveau du parenchyme cérébral, à proximité des neurones dopaminergiques<sup>236</sup> et qu'elles sont capables de produire des cytokines pro-inflammatoires telles que l'IL-6 et l'IL-1 $\beta$  qui pourraient être à l'origine de l'activation de la réponse immunitaire observée dans la MP<sup>237,238</sup>. Ce rôle est étayé par l'identification de polymorphismes dans les gènes *HLA*, des marqueurs de surface exprimés par les cellules présentatrices d'antigènes<sup>23</sup>. L'expression des récepteurs dopaminergiques de types D1 et D2 like à leur surface suggère que ces cellules puissent être sensibles aux variations du niveau de dopamine et présenter des changements moléculaires similaires à ceux des neurones pigmentés de la substance noire<sup>239,235</sup>. En 2007, Scherzer *et al.* ont effectué la première étude du transcriptome de cellules sanguines de cas sporadiques révélant ainsi des variations d'expression de gènes impliqués dans la voie de dégradation ubiquitine-protéasome et dans le transport des récepteurs à la dopamine<sup>240</sup> que l'on sait altéré au niveau cérébral. En 2011, notre laboratoire a également contribué à montrer le potentiel des cellules mononucléées périphériques sanguines par l'analyse du transcriptome de patients porteurs d'une mutation G2019S de *LRRK2* en montrant des perturbations des mécanismes de dégradation tels que le système ubiquitine-protéasome, mais aussi de l'oxydation mitochondriale, de l'inflammation, de la guidance axonale, de la signalisation calcique et de la mort cellulaire par apoptose<sup>241</sup>. D'autres mécanismes ont été mis en lumière tels que la transduction du signal, la signalisation cellulaire, la phosphorylation, la maintenance du cytosquelette, l'endocytose, le transport vésiculaire, le métabolisme des lipides, la guidance axonale et l'infiltration leucocytaire<sup>241</sup>.

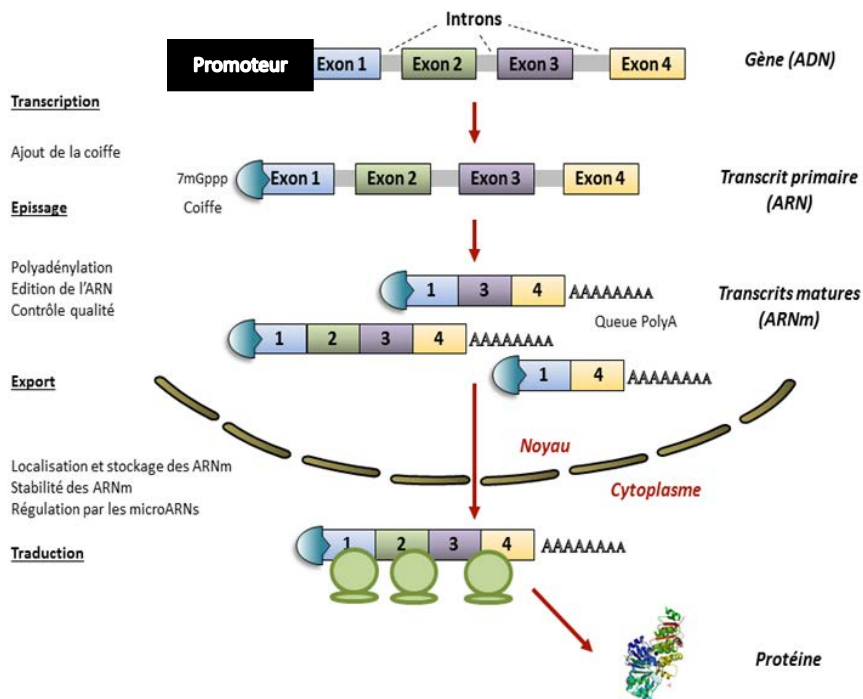
Le recouvrement entre les voies identifiées au niveau cérébral et périphérique<sup>242,243,244</sup> prouve la pertinence de ces cellules pour la compréhension des mécanismes de la maladie et rend alors possible l'identification de modifications précoces pouvant servir de biomarqueurs et/ou de cibles thérapeutiques. Il est à noter que notre étude du transcriptome de cellules mononucléées périphériques sanguines de sujets porteurs de la mutation G2019S de *LRRK2* a révélé des

perturbations de mécanismes connus chez les cas sporadiques montrant que l'étude du transcriptome des formes génétiques pourrait aider à la compréhension des mécanismes qui sous-tendent les formes majoritaires de la maladie. Il est intéressant de noter également que dans cette étude nous avons observé la variation d'expression de gènes impliqués dans la transcription, la maturation des ARNs et la traduction<sup>241</sup>. Ces processus encore peu explorés et que nous avons choisi d'étudier pourraient avoir un rôle crucial dans les processus physiopathologiques de la maladie, ce d'autant que ces dernières années de plus en plus de données de la littérature révèlent l'importance de l'épissage dans les processus physiopathologiques des maladies neurodégénératives.

### 1.1.5. Les perturbations du métabolisme des ARNs dans les maladies neurodégénératives : le rôle de l'épissage

#### 1.1.5.1. La définition de l'épissage

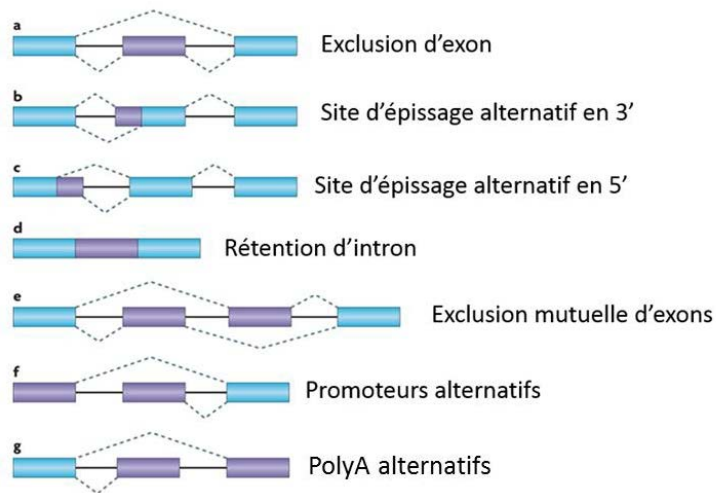
La première étape de l'expression génique est la transcription qui, comme nous l'avons vu, permet la synthèse des transcrits (ARNs) de la cellule. Cette synthèse s'effectue sous la forme de précurseurs appelé transcrits primaires<sup>245</sup> qui se caractérisent par la présence en leur sein de deux types de séquences : les exons et les introns. A la différence des exons (séquences codantes), les introns (séquences non codantes) sont éliminés des transcrits primaires pour générer les transcrits dits matures par un processus appelé épissage. Ce dernier permet l'excision des introns et la ligation des exons. Il définit alors la composition en exons des transcrits matures. On en distingue deux types : l'épissage constitutif où les exons sont systématiquement reconnus et intégrés dans le transcrit mature et l'épissage alternatif où les exons sont soit inclus soit exclus du transcrit mature (figure 18).



**Figure 18 : Les différentes étapes de l'expression génique (d'après <http://www.crcl.fr>)**

La transcription est la première étape de l'expression génique qui va permettre la synthèse, dans le noyau, des transcrits primaires à partir des gènes contenus dans le génome. Dans ce cas précis, la transcription s'effectue par l'ARN polymérase de type II qui va permettre la synthèse de transcrits primaires qui vont subir plusieurs étapes pour leur maturation : l'ajout de la coiffe 7-méthyl-guanosine (7mGpp) en 5', l'épissage qui permet de générer des transcrits matures de composition en exons différents (ici, nous avons 3 transcrits matures du même gène) et l'ajout de la queue polyadénylée (polyA) en 3'. Ces ARNs vont ensuite être exportés vers le cytoplasme où ils seront traduits en protéines, dans le cas des ARNm, ayant des fonctions et/ou des propriétés différentes

L'épissage alternatif est sous l'influence de divers signaux tels que le type cellulaire, le stade de développement, les stimuli externes à la cellule ainsi que son état physiologique ; c'est pourquoi il est décrit comme étant le mécanisme à l'origine de la diversité et de la complexité du génome humain (figure 19). A la découverte de ce processus dans les années 1977, on pensait que seul 5% des gènes humains étaient soumis à l'épissage alternatif. On sait aujourd'hui que plus de 90% des gènes codés par le génome humain y sont soumis faisant de ce mécanisme un processus majeur de l'expression génique<sup>246</sup>. Cette alternance d'inclusion/exclusion permet donc à partir d'un seul et unique transcrit primaire de générer plusieurs transcrits matures codant pour différentes isoformes protéiques ayant des propriétés et des fonctions différentes. C'est ainsi que le génome humain qui dénombre environ 24000 gènes contient plus de 100 000 protéines différentes.



**Figure 19: Diversité des transcrits générés par épissage alternatif (d'après Keren, Lev-Maor, & Ast, 2010)**

Les rectangles bleus et violets représentent différents exons ; les traits pleins noirs représentent les introns et les traits pointillés correspondant aux différents événements d'épissage (clivage) qui sont réalisés dans la cellule. L'épissage de type exclusion d'exon ou cassette-exon est l'épissage le plus fréquent et compte pour environ 40% des événements d'épissage chez les eucaryotes supérieurs ; on retrouve également des événements d'épissage de type 5' et 3' alternatifs où un fragment d'exon soit en 5' soit en 3' va être éliminé en même temps que l'intron avoisinant. Ces types d'épissage représentent respectivement environ 8 et 18% des épissages observés. La rétention d'intron, où l'intron qui normalement devait être épissé est maintenu dans le transcrit mature, constitue l'événement d'épissage le plus rare chez les vertébrés et les invertébrés (moins de 5%). Les épissages de type promoteurs alternatifs ou polyA alternatifs modifient les extrémités non traduites des transcrits où se trouvent souvent des séquences régulatrices jouant sur le devenir de ces transcrits dans la cellule

#### 1.1.5.2. La machinerie effectrice de l'épissage : le spliceosome

La réaction d'épissage est catalysée par un complexe ribonucléoprotéique appelé spliceosome, constitué de 5 sous-unités : les snRNPs (small nuclear ribonucleoprotein particules) U1, U2, U4, U5 et U6. Elles se composent de petits ARNs nucléaires non codants, les snRNA (small nuclear RNA) U1, U2, U4, U5 et U6 en interaction avec les protéines Sm et les protéines Sm-like<sup>151</sup> (figure 20).

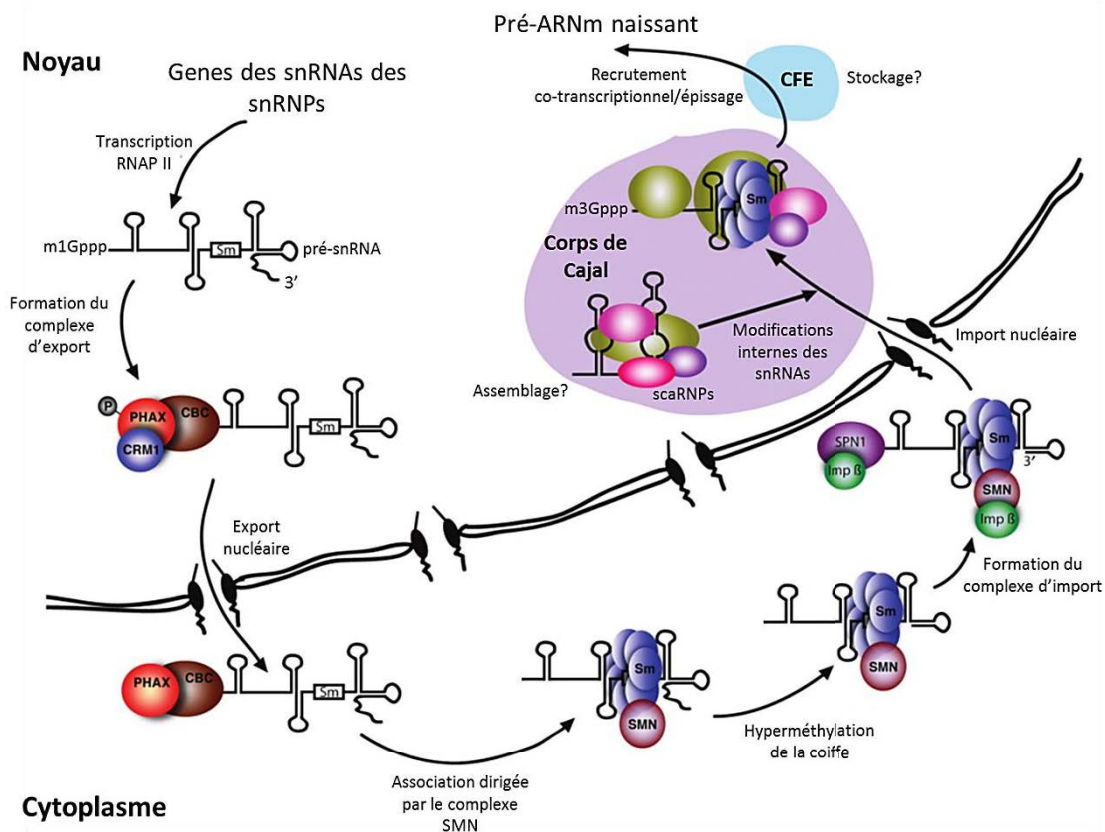
snRNP	U1	U2	U4	U5	U6
	B/B'	B/B'	B/B'	B/B'	Lsm2
	D1	D1	D1	D1	Lsm3
	D2	D2	D2	D2	Lsm4
	D3	D3	D3	D3	Lsm5
	E	E	E	E	Lsm6
	F	F	F	F	Lsm7
	G	G	G	G	Lsm8
	U1-A	U2-A'	15.5K	15K	
	U1-C	U2-B''	61K	40K	
	U1-70K	SF3a		100K	
		SF3b		102K	
		P14		116K	
		SF2/ASF		200K	
		hPrp5p		220K	
		hPrp43			
		SR140		27K <sup>t</sup>	20K <sup>d</sup>
		CHERP		65K <sup>t</sup>	90K <sup>d</sup>
		U2AF		110K <sup>t</sup>	60K <sup>d</sup>
		Hsp60			
		Hsp75			
		PUF60			
		SPF45			
		SPF31			
		BRAF35			

**Figure 20 : Composition des 5 sous-unités du spliceosome (d'après Patel & Bellini, 2008)**

Sept protéines Sm ont été caractérisées ; il s'agit des protéines B/B', D1, D2, D3, E, F et G qui s'assemblent sous forme d'heptamère autour du site Sm AAUUUGUGG conservé entre les différents snRNAs. A la différence des autres snRNAs, le snRNA U6 interagit avec des protéines LSm (Sm-like). La spécificité des snRNP est apportée par l'interaction d'autres protéines non Sm ou non LSm qui sont spécifiques de chaque complexe.

La biogenèse des snRNPs fait intervenir des étapes nucléaires et cytoplasmiques permettant l'assemblage des différentes protéines nécessaires à la réaction d'épissage (figure 21).



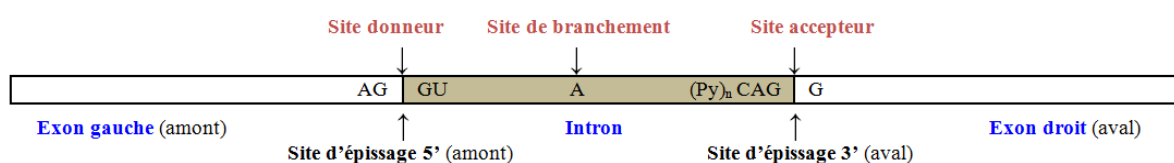


**Figure 21 : Assemblage des snRNPs (d'après Patel & Bellini, 2008)**

Cet assemblage commence dans le noyau dès la fin de la transcription des snRNAs par l'ARN polymérase de type 2 (transcription RNAP II). Comme tous les transcrits RNAP II, les snRNAs contiennent une coiffe en 5' de l'ARN. Cette coiffe est reconnue par les protéines de liaison à la coiffe CBP20 (Cap Binding Protein 20) et CBP80 (Cap Binding Protein 80). Ces protéines vont permettre le recrutement de la protéine PHAX (Phosphorylated Adaptor for RNA export), CRM1 (Chromosomal Maintenance 1) et CBC (Cap-Binding Complex) qui vont permettre l'export nucléaire des snRNAs à travers les complexes de pores nucléaires. Une fois dans le cytoplasme, les protéines Sm vont s'assembler autour des snRNAs pour former les snRNPs. Ceci va être facilité par le complexe SMN (Survival of Motor Neuron protein) composé essentiellement de 7 protéines Gemin (Gemin 2 à 8) qui interagissent à la fois avec les protéines Sm et les snRNAs. La réaction d'épissage ayant lieu dans le noyau, ces snRNPs nouvellement formés sont ensuite adressés à ce compartiment. La méthyltransférase Tgs1 appartenant au complexe SMN va ensuite transférer 2 groupements méthyle à la coiffe en 5' des snRNAs entraînant une hyperméthylation nécessaire au recrutement de la protéine Imp β (Importin β). Imp β constitue un signal de relocalisation nucléaire permettant l'entrée des snRNPs dans le noyau où les snRNPs vont s'orienter vers les corps de Cajal, un sous-compartiment nucléaire dans lequel s'effectue la dernière étape de leur maturation. Bien qu'elle ne soit pas bien connue, il semblerait que les snRNAs y subissent un ensemble d'étapes de maturation incluant une 2'-O-méthylation et une pseudouridylation guidées par les scaRNAs (small Cajal bodies specific RNA). Les snRNPs matures sont ensuite stockés dans des compartiments dits compartiments de facteurs d'épissage (CFE) avant d'être recrutés pour la réaction d'épissage.

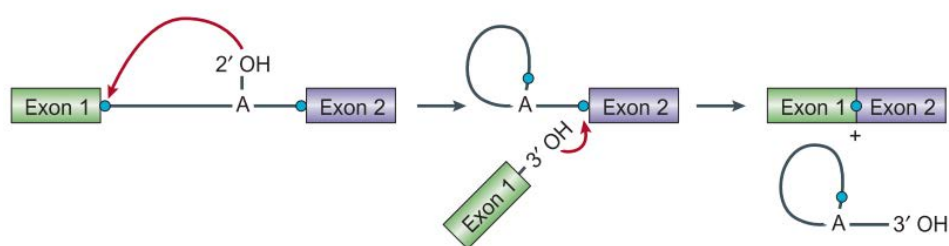
### 1.1.5.3. La réaction d'épissage

Une fois assemblés, les snRNAs contenus dans les snRNPs vont interagir avec les transcrits primaires à épisser par complémentarité de base au niveau de sites spécifiques appelés sites d'épissage en 5' (site donneur) et site d'épissage en 3' (site accepteur). Ces sites délimitent les exons et les introns permettant la reconnaissance des séquences à exciser. Il existe un consensus au niveau des sites d'épissage dont la séquence la plus souvent reconnue est un motif GU/AG<sup>248</sup> (GU étant les deux premiers nucléotides de l'intron, AG les deux derniers, figure 22).



**Figure 22: Motifs des sites d'épissage** (<http://www.cours-pharmacie.com/biologie-moleculaire/transcription-de-ladn.html>)

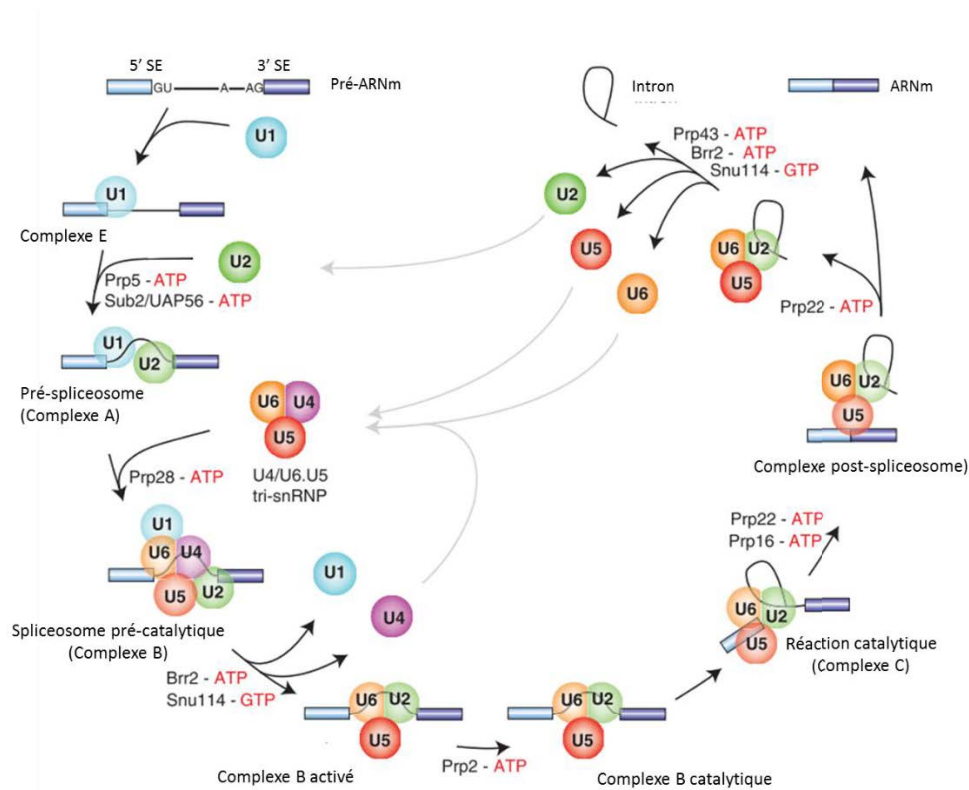
A noter que dans de rares cas ces séquences sont dégénérées et donnent lieu à des sites d'épissage de type AU/AC<sup>248</sup>. Localisé à environ 15-50 nucléotides en amont du site accepteur, on trouve le point de branchement correspondant à un nucléotide adénosine suivi de la queue riche en polypyrimidine (Py)<sub>n</sub> qui vont intervenir dans la réaction d'épissage (figure 24). Cette réaction commence par une attaque nucléophile du groupement 2'OH du point de branchement sur le site donneur en 5' (figure 23).



**Figure 23: Principe de la réaction d'épissage (d'après Patel and Steitz, 2003)**

Ceci permet d'éliminer la liaison de cette extrémité de l'intron avec l'exon en amont (exon 5') et entraîne une liaison de cette extrémité 5' libre avec le point de branchement. Il y a alors formation d'une structure caractéristique appelée lasso ou lariat correspondant à la configuration qu'adopte l'intron une fois son extrémité 5' libérée. Cette réaction est suivie d'une seconde attaque nucléophile du groupement 3'OH de l'exon en 5' sur le site accepteur en 3' qui permet le décrochage de l'intron

et la liaison des exons situés de part et d'autre de l'intron nouvellement excisé. Ces diverses réactions s'effectuent grâce aux snRNPs et aux protéines avec lesquelles elles interagissent qui vont se fixer à l'ARN selon un ordre hiérarchique (figure 24).



**Figure 24 : Réaction d'épissage (d'après Will & Lührmann, 2011)**

Dans un premier temps, snRNP U1 va reconnaître le site d'épissage en 5' par appariement de base via le snRNA qui le compose (snRNA U1). Cette reconnaissance est facilitée par l'ARN polymérase II (qui est en interaction avec le transcrit primaire lors de sa synthèse) dont le domaine C-terminal interagit avec snRNP U1. Les protéines SF1/mBBP et U2AF65 (sous-unité complexe U2) viennent ensuite se fixer au site de branchement et à la queue polypyrimidine respectivement. Le site en 3' est lui reconnu par la protéine U2AF35 formant ainsi le complexe E du spliceosome. Par l'intervention des hélicases Prp5 et Sub2, snRNP U2 va reconnaître des séquences au niveau du site de branchement, rapprochant snRNP U2 de snRNP U1 et permettant leur interaction. C'est la formation du complexe A pré-spliceosome. Après l'assemblage du complexe A, le complexe tri-snRNPs U4/U6 et U5 est recruté au niveau du site en 5' pour former le complexe B pré-catalytique. Ce recrutement est facilité par l'hélicase Prp28 et génère une série de remaniement qui va permettre la libération de snRNP U1 et snRNP U4 pour ainsi former le complexe B activé. Ce complexe activé va catalyser la réaction d'épissage et induire la formation du complexe post-spliceosome qui marque la fin de la réaction d'épissage. Les snRNPs U2, U5 et U6 vont alors se décrocher et être recyclés pour la prochaine réaction

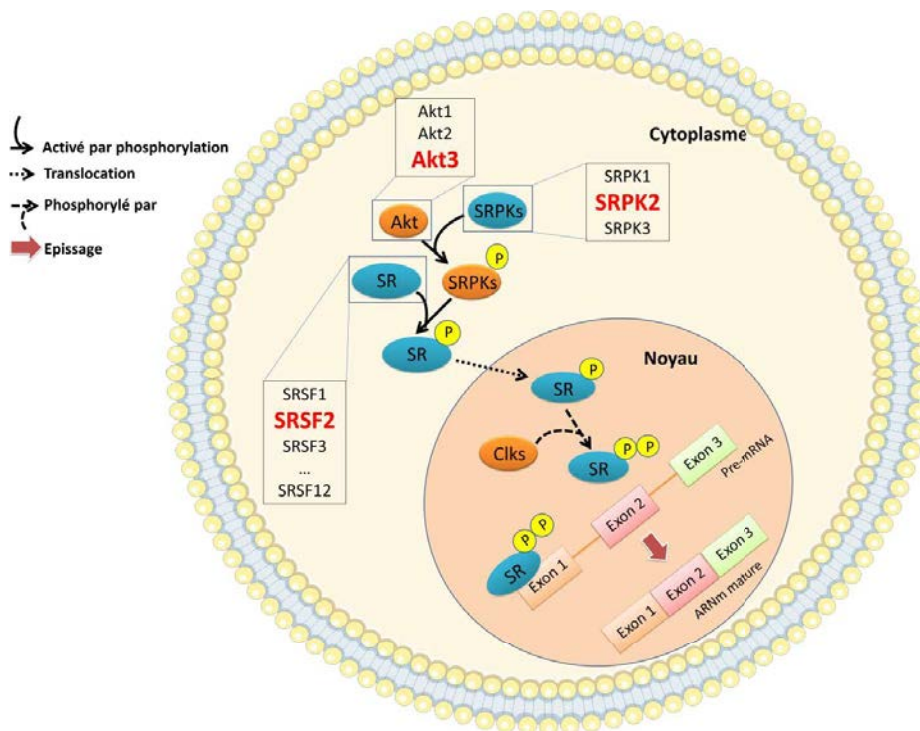
#### 1.1.5.4. La régulation de l'épissage

Plusieurs mécanismes de régulation assurent la fidélité de la réaction d'épissage. Ils garantissent une bonne reconnaissance des sites et une fixation stable des facteurs d'épissage sur l'ARN. Ces mécanismes font intervenir des facteurs en *cis* et en *trans* incluant les séquences *cis*

régulatrices, les protéines SR et hnRNPs, la structure de l'ARN et de la chromatine, l'état transcriptionnel et les ARNs non codants.

Les séquences cis-régulatrices de longueurs et de séquences variables, sont réparties tout au long de l'ARN. On distingue les séquences activatrices et les inhibitrices nommées ESE (exonic splicing enhancers) lorsqu'elles sont localisées dans un exon et ISE (intronic splicing enhancers) lorsqu'elles se trouvent dans les introns. De la même manière, les séquences inhibitrices présentes dans les exons sont appelées ESS (exonic splicing silencers) et ISS (intronic splicing silencers) lorsqu'elles sont présentes dans les introns<sup>250</sup>. Ces séquences sont reconnues par des éléments en *trans* que sont les protéines SR (riches en sérines et en arginines) et les protéines hnRNPs respectivement. Elles appartiennent toutes deux à la famille des protéines de liaison aux ARNs (protéines RBP pour RNA Binding Proteins) par l'intermédiaire de domaines d'interaction spécifiques comme les motifs RRM (RNA recognition motif). Le rôle des protéines SR dans l'épissage est déterminant car elles permettent le recrutement des sous-unités du spliceosome au niveau de l'ARN prématuré<sup>251,250,252,253</sup>. C'est le cas par exemple de la protéine SR SF2/ASF qui permet le recrutement de snRNP U1 au niveau du site d'épissage en 5'<sup>251,250,252</sup>. A la différence des protéines SR, le mode d'action des protéines hnRNPs est moins bien connu. Toutefois, il semblerait que, par interaction avec d'autres protéines, elles entraînent une multimérisation le long des exons bloquant ainsi le recrutement des snRNP.

La régulation de l'épissage fait également appel à des modifications post-traductionnelles des facteurs d'épissage telles que la phosphorylation des protéines SR. En effet, ces protéines possèdent de nombreux acides aminés phosphorylables qui sont la cible des protéines kinases de la famille des SRPKs, CLK/Sty, DYRKs<sup>254,255,256</sup> ou encore Akt<sup>255,257</sup> (figure 25).



**Figure 25: Régulation de l'épissage par phosphorylation des protéines SR**

La phosphorylation des protéines SR contrôle leur localisation dans la cellule. Présentes dans le cytoplasme sous la forme hypophosphorylée, leur hyperphosphorylation par les différentes kinases entraîne leur translocation dans le noyau où elles peuvent réguler l'épissage. Les kinases des protéines SR peuvent également se phosphoryler entre elles ; c'est le cas de la phosphorylation des protéines SRPKs par Akt qui module leur activité kinase.

L'activation des kinases est modulée par différents signaux extracellulaires ; c'est le cas par exemple des facteurs de croissance ou de l'insuline qui modulent l'état d'activation de la voie de signalisation PI3K/Akt. Cette phosphorylation est inversée par de nombreuses phosphatases telles que PP1 (Protein phosphatase 1) et PP2A (Protein phosphatase 2A), l'état de phosphorylation des protéines SR ayant des répercussions au niveau de l'épissage en jouant sur la localisation nucléaire de ces facteurs<sup>258,259,260</sup>. Des études immunologiques ont permis de localiser les facteurs d'épissage au niveau de différents sous-compartiments nucléaires non membranaires tels que les corps de Cajal, les gems ou encore les speckles nucléaires. Il semblerait que leur localisation au niveau des corps de Cajal et les gems soit importante pour la maturation des snRNPs (paragraphe 1.1.5.2), les speckles étant considéré comme un lieu de stockage dans l'attente du recrutement de ces facteurs pour la réaction d'épissage qui s'effectue dans le nucléoplasme<sup>226, 264,234-236</sup>.

L'ARN à épisser lui-même peut intervenir dans la régulation de son épissage en adoptant une structure secondaire pouvant affecter la reconnaissance des sites d'épissage par les éléments du spliceosome<sup>265,266,267</sup>. C'est pour cela que le spliceosome compte parmi ces protéines un grand

nombre de DExD/H RNA hélicases qui déroulent la structure de l'ARN et influent sur les interactions ARN-protéines nécessaires à la réaction d'épissage. D'autres ARNs, comme les miARNs et les ARNs longs non codants peuvent intervenir dans la régulation de l'épissage en modulant l'expression de ces facteurs ou en les séquestrant, les rendant peu disponibles pour la réaction d'épissage. C'est le cas des miARNs miR-133 et miR-23a/b dont la surexpression induit une diminution de l'expression de facteurs d'épissage de la protéine hnRNP I<sup>268,269</sup> ou de l'ARN long non codant MALAT 1 localisé dans les speckles où il interagit avec la protéine SR SF2/ASF et la maintien dans ce sous-compartiment<sup>270</sup>.

Ainsi, l'épissage touchant la quasi-totalité des gènes exprimés chez l'homme, il n'est pas surprenant que des perturbations de ce processus, que ce soit au niveau de ses effecteurs ou de ses régulateurs, puissent contribuer ou être à l'origine de maladies humaines<sup>271</sup>.

#### 1.1.5.5. Les maladies neurodégénératives : des exemples de perturbations de l'épissage

Ces dernières années, la littérature scientifique a rapporté de nombreuses perturbations de l'épissage dans les maladies neurodégénératives ; l'atrophie musculaire spinale, la dystrophie myotonique ou encore la sclérose latérale amyotrophique en sont quelques exemples.

L'atrophie musculaire spinale est une maladie neuromusculaire qui est associée à la dégénérescence et à la perte de neurones de la corne antérieure de la moelle épinière. Cette dégénérescence conduit à une perte musculaire progressive et dans les cas les plus sévères à une détresse respiratoire suivie de la mort du patient. L'origine de cette maladie réside dans des mutations ou des délétions des gènes *SMN1* et *SMN2* qui codent des protéines SMN (survival motor neuron) du complexe d'assemblage des snRNPs. La déficience de ces protéines SMN a été montrée comme étant à l'origine d'une diminution du nombre de snRNPs matures suggérant des défauts d'épissage qui sont délétères pour les modèles animaux de la maladie<sup>272</sup>. Des modifications de l'épissage ont également été observées dans la dystrophie myotonique, une maladie liée à des répétitions de motifs nucléotidiques dans les gènes *DMPK* (dystrophia myotonica-protein kinase) ou *ZNF9* (Zinc Finger Protein 9) favorisant l'agrégation de leurs ARNs (foci) dans le noyau de la cellule. Ces foci engendrent des changements d'expression et la séquestration de régulateurs de l'épissage tels que les protéines MBNLs, CUGBP1 et hnRNP H réduisant ainsi leur disponibilité et modifiant le profil d'épissage de la cellule. La séquestration de ce type de protéines est également observée dans la sclérose latérale amyotrophique (SLA), une maladie qui bien qu'elle soit essentiellement sporadique, est liée dans de rares cas à des mutations de nombreux gènes codant des protéines hnRNPs telles que TDP-43, FUS, EWSR1, TAF15, hnRNPA1 ou encore hnRNPA2/B1. Cette

séquestration est assurée par des foci formés par les ARNs du gène *C9orf72* au sein duquel des expansions du motif nucléotidiques GGGCC sont liées à la maladie mais aussi par les expansions de polyglutamine dans l'ataxine-2 qui séquestre la protéine TDP-43.

Si les mutations dans le gène de l'ataxine-2 sont un facteur de risque de la SLA, elles sont aussi, comme nous l'avons mentionné précédemment, responsables de la MP. Connaissant le rôle de l'ataxine-2 dans le métabolisme des ARNs (paragraphe 1.1.4.2.3.) et aux vues des variations d'expression génique liées à ce métabolisme dans les cellules mononucléées périphériques sanguines de parkinsoniens porteurs de la mutation G2019S de *LRRK2*, ces données ont renforcé notre intérêt pour l'étude de l'épissage dans la MP. Reste à définir si de telles variations sont pertinentes (existantes réellement) et constituent un mécanisme physiopathologique de la maladie. L'effet de la surexpression de protéine *LRRK2* mutée G2019S dans des lignées cellulaires de neuroblastome SK-N-MC sur l'épissage de l'exon 1 de *MAPT* suggère que cette protéine peut intervenir directement ou indirectement dans ce processus de maturation des ARNs<sup>273</sup>. L'expression préférentielle de certains transcrits de gènes liés aux formes monogéniques (*PRKN*, *LRRK2* et *SNCA*) chez les sujets parkinsoniens par rapport aux témoins ou dans des lignées cellulaires dopaminergiques de neuroblastome SY5Y exposées au MPP<sup>+</sup>, à la roténone, à la 6-OHDA implique que certains transcrits sont associés à la maladie<sup>274,275</sup>. La surexpression de l'isoforme 112 de l' $\alpha$ -synucléine, plus prompt à l'agrégation et à la formation des corps de Lewy que l'isoforme 140<sup>276</sup>, est associée à une diminution de l'activité du protéasome dans des lignées cellulaires et les souris exposées au MPP<sup>+</sup>, incluant une inhibition de l'activité de la chymotrypsine like (une protéase du protéasome)<sup>275</sup>. Ces données supposent que l'épissage pourrait participer à la genèse de la maladie, une hypothèse reprise par Alieva *et al.* qui suggère que ces dérèglements surviennent dès les premiers stades de la MP contribuant ainsi à sa genèse<sup>277</sup>.

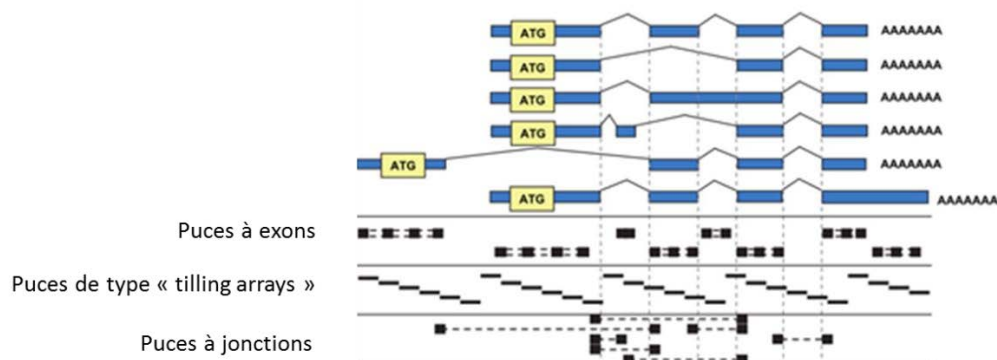
Une analyse plus approfondie est nécessaire pour mesurer l'importance et l'étendue des défauts d'épissage dans la physiopathologie de la maladie. Le développement de nouvelles techniques d'étude du transcriptome rendent désormais cette analyse possible.

#### 1.1.5.6. Les méthodes d'étude de l'épissage

Comme nous venons de le voir, l'épissage fait intervenir de nombreux ARNs et protéines partenaires qui vont définir la composition en exons des transcrits de la cellule. Plusieurs techniques ont été développées afin d'étudier ces événements d'épissage (type de jonctions) et connaître les transcrits exprimés par la cellule. Elles reposent soit sur l'utilisation de sondes oligonucléotides complémentaires soit sur le séquençage des transcrits.

### 1.1.5.6.1. Les puces d'épissage

On distingue 3 types de puces d'épissage différentes : les puces à exons, les puces à jonctions et les puces de type tiling arrays (figure 26).



**Figure 26 : Analyse des événements d'épissage par l'utilisation de micropuces (Kolle, 2007)**

Les rectangles bleus représentent les différents exons présents dans chaque transcrit et séparés par les introns. Les rectangles noirs correspondent aux sondes des différents types de puces d'épissage existantes qui sont soit à cheval sur les exons (puces à exons), soit à cheval sur les jonctions (puces à jonctions) ou possèdent une association de ces deux types de sondes (puces de type tiling arrays)

Comme les puces d'expression (paragraphe 1.1.4.5.), ces puces d'épissage reposent sur l'utilisation de sondes oligonucléotides fixées à un support et capables de s'apparier à différentes régions des transcrits par complémentarité de base et de mesurer des niveaux d'expression relative. Selon la nature de la sonde, ces puces permettent de mesurer les niveaux d'expressions des exons (puces à exons), des jonctions entre exons (puces à jonctions) ou les deux (puces de type tiling array).

Les puces à exons sont les premières puces à avoir été développées. Elles possèdent des sondes qui couvrent l'ensemble des exons présents dans les bases de données. Ainsi ces sondes permettent de détecter la présence de tel ou tel exon dans un échantillon. Les puces à exons les plus utilisées sont les puces Affymetrix all exon qui existent pour l'étude du génome humain, souris et rat. Cependant ces puces présentent certaines limites :

- La longueur des sondes (de 25 à 60 nucléotides) qui requiert que la séquence de l'exon soit d'au moins 25pb ; les exons de longueur inférieure ne sont pas identifiés.
- Du fait de l'épissage alternatif, un exon peut être présent dans plusieurs transcrits et les puces à exons sont incapables d'identifier l'origine de l'exon donc de savoir de quel transcrit il provient



Les puces à jonctions ont été développées par la suite. Elles reposent également sur l'utilisation de sondes conçues d'après les données de séquence des bases de données mais à la différence des puces à exons, ces sondes sont complémentaires des jonctions entre exons. Elles apportent une information complémentaire à celle des puces à exons car elles permettent non seulement d'identifier les exons exprimés dans un échantillon mais également de connaître les exons liés entre eux grâce aux sondes à cheval sur les jonctions. Cependant, elles ne permettent pas de reconstituer les transcrits.

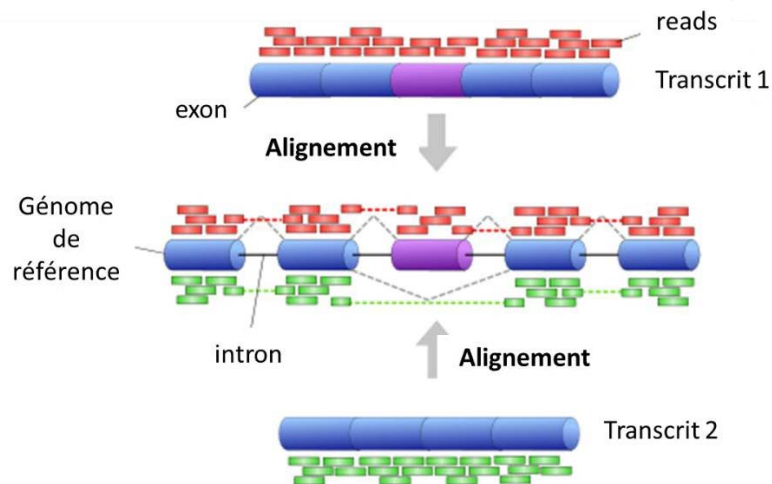
Enfin les puces de type « tiling arrays » présentent l'avantage de combiner à la fois des sondes présentes sur les exons et celles présentes sur les jonctions. Ainsi, elles fournissent une information supplémentaire qui permet de déterminer si un exon provient d'un transcrit particulier ou s'il est commun à plusieurs transcrits.

Bien que ces puces soient informatives, comme toute technique, elles présentent certaines limites. L'une d'entre elles est qu'elles reposent exclusivement sur l'utilisation de sondes conçues à partir des données de séquence identifiées dans les bases de données. La découverte de nouvelles jonctions non encore décrites dans ces bases est alors impossible. De ce fait, ces puces ne permettent pas la mise en évidence de nouvelles jonctions entre exons et de nouveaux transcrits, qui pourraient être générés suite à des défauts d'épissage. De plus, comme toutes les puces à oligonucléotides, la possibilité d'hybridations croisées aspécifiques existe et génère des faux positifs.

#### 1.1.5.6.2. Le séquençage des ARNs

Le développement des méthodes de séquençage nouvelles générations et des outils d'analyse à haut débit a contribué à l'amélioration des techniques d'étude du transcriptome et de l'épissage en particulier. Ces techniques reposent sur les méthodes de séquençage de nouvelles générations qui, contrairement aux méthodes de séquençage classiques de type Sanger, produisent rapidement et simultanément un grand nombre de séquences. Elles permettent de déterminer la séquence de l'ensemble des transcrits présents dans un échantillon, ce qui contrairement aux puces d'épissage, donne l'accès à l'ensemble des transcrits présents dans un échantillon et pas seulement à ceux qui s'apparient sur les sondes. Le séquençage élargit le champ des transcrits étudiés favorisant l'identification de tous les événements d'épissage de la cellule (connus ou non des bases de données). Les capacités des séquenceurs actuels ne permettant pas de séquencer en une seule fois l'entièreté des transcrits, ces derniers doivent être clivés en fragments plus petits qui vont être séquencés. De plus, ces séquenceurs étant des séquenceurs à ADN, les ARNs doivent être convertis en ADN complémentaires (ADNc ayant la même séquence que l'ARN). Ainsi le séquençage des ARNs

s'effectue suivant plusieurs étapes, que nous verrons plus en détail par la suite et qui incluent : la récupération des ARNs, la fragmentation des ARNs et la synthèse d'ADNc complémentaires qui vont être alignés sur un génome de référence. Cet alignement permet de déterminer les transcrits dont ces fragments sont issus et d'avoir une information sur leur séquence (enchâînement d'exon et/introns retenus, figure 27).



### Figure 27 : Principe du RNAseq

*Les ARNs totaux sont extraits à partir d'un échantillon puis enrichis en ARNm par déplétion des ARNr majoritaires dans la cellule. Les ARNs sont ensuite fragmentés puis soumis à une étape de transcription inverse qui permet la synthèse de fragments d'ADNc dont les extrémités vont être séquencées. Ces fragments séquencés appelés « reads » sont alignés sur un génome de référence. Cet alignement conduit à l'identification des gènes ou transcrits dont sont issus les reads grâce aux annotations présentes dans le génome de référence utilisé.*

En plus de permettre l'identification des transcrits et de leur composition en exons, le RNAseq donne accès à des informations sur la quantité des transcrits exprimés fournissant une mesure quantitative et une analyse d'expression différentielle par comparaison de différents échantillons. Le RNAseq apporte donc des notions qualitatives et quantitatives sur les profils d'expression génique. Il présente de nombreux autres avantages par rapport aux techniques d'étude du transcriptome reposant sur l'utilisation de sondes complémentaires (tableau 2) qui font de cette technique l'outil idéal pour une analyse approfondie du transcriptome.

	<b>PUCES</b>	<b>RNAseq</b>
<b>Génome de référence</b>	Oui, pour la création des sondes	Oui, pour l'alignement des reads mais pas pour le séquençage <i>de novo</i>
<b>Technique</b>	Hybridation par complémentarité de base avec les sondes	Accès direct à la séquence
<b>Système de détection</b>	Lecture de la fluorescence par un scanner: problème de détection dans les faibles et les fortes intensités	Précision de la lecture à une base près permettant l'étude des SNP
<b>Coût</b>	+	+++
<b>Reproductibilité</b>	+++++	+++++
<b>Obtention des résultats</b>	Rapide	Plus longue
<b>Quantité de données générées</b>	+	+++
<b>Outils d'analyse statistiques</b>	Bien définis	Encore en développement
<b>Analyses</b>	Seulement des jonctions connues	Analyse des jonctions connues, découverte de nouvelles jonctions et transcrits, quantification de l'expression des gènes et transcrits

**Tableau 2: Synthèse des avantages du RNAseq par rapport aux puces d'épissage**

Notre hypothèse de l'existence de perturbations de l'épissage dans la MP, basée sur nos données transcriptomiques de sujets porteurs de la mutation G2019S de *LRRK2*, pourrait alors être confirmée par l'utilisation d'une telle technique offrant une exploration plus approfondie du transcriptome. Toutefois, cette hypothèse étant basée sur des données obtenues à partir d'un sous groupe de parkinsoniens (mutation G2019S de *LRRK2*), nous ne pouvons pas exclure que ces perturbations soient propres à ces sujets et que les perturbations de l'épissage ne soient pas un mécanisme généralisé à l'ensemble des cas. Une analyse comparée de différentes formes de la maladie est donc nécessaire afin de déterminer si les perturbations observées chez ces sujets

reflètent des modifications observées chez un grand nombre de malades avant de proposer l'épissage comme mécanisme physiopathologique de la maladie.

# OBJECTIFS

## 2. Objectifs de la thèse

La compréhension des mécanismes moléculaires de la MP est un enjeu majeur de la recherche actuelle. L'étude des formes génétiques de la maladie, par des approches de génétiques classiques, a montré leur potentiel dans l'acquisition des connaissances des mécanismes de la maladie. Cependant, face à l'étendue des perturbations moléculaires qui sous-tendent la maladie, l'utilisation de technologies apportant des informations à grande échelle est inéluctable. C'est pourquoi notre laboratoire a choisi de recourir à l'analyse du transcriptome de cellules mononucléées périphériques sanguines dérivées de patients afin d'approfondir les connaissances des mécanismes moléculaires de la MP. Les données préliminaires du laboratoire, obtenues avant mon arrivée dans l'équipe du Dr. Marie-Christine Chartier-Harlin, avaient apporté la preuve du concept de l'utilisation de ces cellules comme modèle de la MP. Elles avaient également révélé l'existence de variations d'expression de gènes du métabolisme des ARNs dans les cellules sanguines de patients porteurs de la mutation G2019S de *LRRK2*<sup>241</sup>.

Basée sur ces données, nous avons émis l'hypothèse de l'existence de perturbations de l'épissage dans la MP qui pourraient, si elles se confirment, servir de cibles thérapeutiques et/ou de biomarqueurs précoces. Cependant, cette hypothèse suppose que les variations du transcriptome observées dans les cellules de patients porteurs de la mutation G2019S de *LRRK2* reflètent des modifications observées chez la majorité des malades. Ceci n'ayant pas été vérifié jusqu'alors, nous avons dans un premier temps, effectué l'étude du transcriptome de cellules sanguines de patients porteurs de la mutation G2019S de *LRRK2* et de sujets sporadiques à l'aide de puces d'expression ; ceci dans le but de déterminer si ces sujets présentent des altérations communes et si les défauts du métabolisme des ARNs peuvent être un mécanisme physiopathologique attribuable à différentes formes de la maladie.

En parallèle, nous avons étudié le profil d'expression génique de cellules mononucléées périphériques sanguines de sujets porteurs de mutations du gène *ATXN2*, du fait du rôle connu de sa protéine dans le métabolisme des ARNs. Cette étude devrait nous permettre de savoir si les mutations de ce gène entraînent des changements d'expression de gènes liés au métabolisme des ARNs et d'évaluer le potentiel de ces perturbations dans la genèse de la MP.

Enfin, dans le but de déterminer si ces variations du métabolisme des ARNs incluent des variations quantitatives et qualitatives de l'épissage c'est-à-dire des modifications d'expression et des jonctions entre exons, nous avons choisi d'effectuer une analyse par utilisation de la méthode de séquençage des ARNs. Cette étude devrait nous permettre d'identifier, s'ils existent, des

changements de composition en exons des transcrits et d'évaluer l'importance des ces modifications dans la physiopathologie de la maladie.

# RÉSULTATS



### 3. Résultats

#### 3.1. **Chapitre 1** : Étude du transcriptome de cellules mononucléées périphériques sanguines de sujets porteurs de la mutation G2019S de *LRRK2* et de sujets sporadiques.

La mutation G2019S de *LRRK2* a été décrite comme la mutation causale de la MP la plus commune. Elle conduit à l'apparition d'un phénotype similaire et/ou proche de celui des formes sporadiques notamment au niveau de l'âge d'apparition des premiers symptômes et de la progression de la maladie. Ces similarités ont amené à penser que les mécanismes moléculaires conduisant à leur apparition étaient communs et, du fait de la prépondérance des formes sporadiques et de la forte fréquence des formes de MP liées à la mutation G2019S de *LRRK2*, que ces mécanismes pourraient être responsables de la majorité des cas de MP. Une première étude du transcriptome de cellules mononucléées périphériques sanguines de sujets porteurs de la mutation de *LRRK2* avait montré que cette forme de la maladie était associée à des changements d'expression de gènes impliqués dans les mécanismes de dégradation cellulaire, dans la guidance axonale, dans la mitochondrie, l'inflammation et de l'apoptose<sup>241</sup>. L'observation de ces changements également retrouvés au niveau cérébral avait révélé que l'étude du transcriptome de cellules sanguines reflétait des perturbations des voies de signalisation ayant lieu dans le cerveau. L'identification de certains mécanismes pathologiques à partir de ces cellules est alors envisageable.

Sur la base de cette constatation et afin de déterminer si le développement de la MP repose sur les mêmes mécanismes chez les sujets porteurs de la mutation G2019S de *LRRK2* et les sujets sporadiques, nous avons étudié le transcriptome de cellules mononucléées périphériques sanguines de 20 cas sporadiques, 9 sujets porteurs de la mutation G2019S de *LRRK2* et 40 sujets témoins à l'aide de puces à ADN. Cette analyse nous a permis de comparer les niveaux d'expression génique entre les cas sporadiques *versus* les témoins et les malades porteurs de la mutation de *LRRK2* *versus* les témoins. Ces gènes différentiellement exprimés ont été soumis au logiciel d'analyse Ingenuity Pathways Analysis (IPA®) qui permet de mesurer l'association entre les gènes soumis au logiciel et l'ensemble des gènes connus et admis comme étant impliqués dans un processus cellulaire donné ou voie canonique. En utilisant un test statistique exact de Fischer, IPA® détermine si cette association est due au hasard. Le seuil de probabilité ( $p$ ) que nous avons utilisé pour considérer qu'une association n'était pas due au hasard est  $p < 0,05$ , utilisé de manière empirique.

L'analyse des voies canoniques enrichies en gènes différentiellement exprimés chez les sujets sporadiques par rapport aux témoins a mis en évidence des dérèglements de voies déjà montrés comme étant perturbés chez les parkinsoniens. En effet nous avons retrouvé des modifications des

voies liées à la réponse inflammatoire, la réponse immunitaire, la prolifération et la mort cellulaire. Les perturbations des voies liées à l'inflammation et aux processus immunitaires sont également retrouvés chez les sujets porteurs de la mutation G2019S de *LRRK2* chez qui des altérations des processus de survie cellulaires, des mécanismes du cancer (prolifération cellulaire) et de dégradation cellulaire ont aussi été observées. Afin d'évaluer le potentiel des études transcriptomiques effectuées à partir des PBMC, dans l'identification de biomarqueurs et de cibles thérapeutiques potentiels, nous avons comparé nos données de puces d'expression aux données de la littérature de Scherzer *et al.*<sup>240</sup> et de Shehadeh *et al.*<sup>278</sup> obtenues à partir de sang total. *ALAS2*, *EPB42*, *SELENBP1* et *SLC4A1* sont 4 gènes liés au métabolisme du fer qui sont différentiellement exprimés entre les malades et les témoins dans chacune de ces 3 études et qui du fait de leur variation d'expression répétée dans différents groupes de sujets pourraient être de bons candidats pour distinguer les malades des témoins.

En s'intéressant de plus près aux voies canoniques dérégulées à la fois chez les cas sporadiques et les porteurs de la mutation G2019S de *LRRK2*, nous avons observé des perturbations communes des voies de survie cellulaire (« voie de signalisation PI3K/Akt ») et de la mitochondrie montrant que ces deux groupes de patients partagent des modifications communes déjà décrites comme étant associées à la maladie. L'identification de voies déjà décrites a conforté l'idée du potentiel du transcriptome de ces cellules dans la mise en évidence de nouveaux mécanismes, communs à ces deux formes de la maladie mais aussi à la majorité des cas de MP. De façon intéressante, la voie de signalisation EIF2 qui jusqu'ici n'a pas été fortement explorée, apparaît comme la voie commune la plus significativement perturbée, à la fois chez les sujets sporadiques et les sujets porteurs de la mutation G2019S de *LRRK2*. Pour poursuivre notre analyse, nous avons ré-analysé les données transcriptomiques des équipes de Scherzer<sup>240</sup> et de Shehadeh<sup>278</sup> à l'aide de l'outil GEO2R. Ce dernier nous a permis d'identifier les gènes différentiellement exprimés qui ont été soumis à IPA®. Les voies canoniques les plus retrouvées à la fois dans ces 2 études et dans la notre étaient l'entrée des virus via la voie de l'endocytose, la signalisation de l'infiltration leucocytaire et une nouvelle fois la voie de signalisation EIF2 (Tableau 3).

	G2019S <i>LRRK2</i>	Sporadic PD	Scherzer <i>et al.</i>	Shehadeh <i>et al.</i>	Simunovic <i>et al.</i>	Elstner <i>et al.</i>
Voie de signalisation EIF2	5,01E-17	2,50E-3	3,5 E-2	1,25E-4	4,00E-4	1,10E-3
Entrée des virus via la voie de l'endocytose		2,30E-2			4,00E-4	2,50E-4
Signalisation de l'infiltration leucocytaire	2,10E-2		3,90E-3	1,66E-3	2,50E-4	4,00E-2
Dysfonctions mitochondriales	3,20E-5	4,40E-2		3,72E-3		
Voie de signalisation PI3K/Akt	1,20E-2	4,70E-2		6,03E-6	3,00E-4	1,80E-3
Voie de signalisation du cancer de la prostate	1,40E-2	4,10E-3		3,71E-3		
Voie de signalisation de la leucémie myéloïde	3,80E-2	2,80E-2		2,57E-4		

**Tableau 3: Sélection de voies canoniques IPA® associées aux gènes différentiellement exprimés dans nos données et dans la littérature (p<0,05)**

La MP ayant longtemps été considérée comme une maladie du système nerveux central et touchant majoritairement les neurones dopaminergiques de la substance noire, Simunovic *et al.*<sup>227</sup> et Elstner *et al.*<sup>279</sup> ont analysé le transcriptome de ces cellules obtenues par microdissection laser pour mieux comprendre les changements moléculaires liés à la maladie. Par le biais de ces deux analyses, plusieurs voies canoniques ont pu être identifiées montrant une nouvelle fois des voies associées à la maladie telles que les dysfonctions mitochondriales et la voie d'ubiquitination des protéines mais aussi des perturbations de la voie de signalisation EIF2.

Les similitudes entre nos données et celles obtenues de l'analyse du transcriptome des cellules neuronales de parkinsoniens renforcent l'idée que 1) les perturbations liées à la maladie au niveau cérébral puissent être retrouvées au niveau périphérique, 2) que ces cellules puissent être utilisées comme modèle d'étude afin d'améliorer les connaissances de la maladie. Jusqu'ici, les nombreuses recherches effectuées dans le cadre de l'étude de la MP ont mis en avant les perturbations des processus inflammatoires, des dysfonctions au niveau mitochondrial, au niveau

des mécanismes de dégradation protéique et du système immunitaire mais n'avaient pas mis l'accent sur la voie de signalisation EIF2. Telle que décrite dans IPA®, cette dernière est principalement connue pour son implication dans la régulation de la traduction. Elle joue aussi un rôle dans l'épissage car elle comprend des gènes *RP* (Ribosomal Protein) codant des protéines ribosomiques, qui sont capables de s'auto-épisser (Tableau 4).

Gènes	Ratio (G2019S <i>LRRK2</i> /Témoins)	Ratio (Sporadiques/Témoins)
RPL19	1,35	1,40
RPL21	-1,43/1,32	-1,33/-1,27/-1,32
RPL27A	1,20	1,57
RPL4	1,62	1,60/1,50
RPL9	-1,32	1,40/1,56
RPS2	1,22/1,30	1,34
RPS9	1,30	1,32

**Tableau 4: Sélection de gènes de la voie EIF2 dérégulés chez les parkinsoniens porteurs de la mutation G2019S de *LRRK2* et chez les sujets sporadiques (t-test de Welsch ; p<0,05)**

*Les gènes de la voie EIF2 communs à sujets porteurs de la mutation G2019S de *LRRK2* et aux cas sporadiques sont rapportés dans le tableau. Les différents ratios d'expression indiqués pour un même gène correspondent aux différentes sondes de ce gène qui sont significativement dérégulées*

La mise en évidence des dérèglements de cette voie, encore peu rapportée dans la littérature, laisse penser que nous sommes face à un nouveau mécanisme physiopathologique impliquant la traduction et/ou l'épissage.

[Article 1: Involvement of the immune system, endocytosis and EIF2 signaling in both genetically determined and sporadic forms of Parkinson's disease](#)

## **Involvement of the immune system, endocytosis and EIF2 signaling in both genetically determined and sporadic forms of Parkinson's disease**

Eugénie Mutez, MD, PhD<sup>1,2,3</sup>; Aurore Nkiliza, MSc<sup>1,2</sup>; Karim Belarbi, PhD<sup>1,2</sup>; Amélie De Broucker, MSc<sup>1,2</sup>; Christel Vanbesien-Mailliot, PhD<sup>1,2</sup>; Séverine Bleuse, PhD<sup>1,2,3</sup>; Aurélie Duflot, BSc<sup>1,2</sup>; Thomas Comptdaer, BSc<sup>1,2</sup>; Pierre Semaille, MSc<sup>1,2</sup>; Renaud Blervaque, BSc<sup>4</sup>; David Hot, PhD<sup>4</sup>; Frederic Leprêtre, PhD<sup>5</sup>; Martin Figeac, PhD<sup>5</sup>; Alain Destée, MD, PhD<sup>1,2,3</sup> and Marie-Christine Chartier-Harlin, PhD<sup>1,2\*</sup>

<sup>1</sup> UMR 837 INSERM, Team 6, JPArc, IRCL, Lille, France;

<sup>2</sup> Univ Lille Nord de France, Lille, France;

<sup>3</sup> Movement Disorders Department, Neurologic Clinic, University Hospital, Lille, France;

<sup>4</sup> Center for Infection and Immunity, Institut Pasteur de Lille, Lille, France;

<sup>5</sup> Functional Genomic Platform, UDSL, IFR114, IRCL, Lille, France.

\*Correspondence to: Marie-Christine Chartier-Harlin, UMR 837 INSERM, Team 6, JPArc, IRCL, Place de Verdun, F-59045 Lille cedex, France. Tel: +33 320 169 222, Fax: +33 320 169 229, E-mail: [marie-christine.chartier-harlin@inserm.fr](mailto:marie-christine.chartier-harlin@inserm.fr)

**Keywords:** Parkinson's disease; transcriptome; peripheral blood mononuclear cell; PBMC; LRRK2 mutation; immune response; EIF2 signaling; endocytosis

## **Abstract**

The leucine-rich repeat kinase 2 (LRRK2) G2019S mutation is a common genetic cause of Parkinson's disease (PD). Although patients with sporadic PD and individuals with LRRK2-linked PD display the classical PD phenotype, it is not known whether or not the same biological pathways are deregulated in each context. By using transcriptome profiling, we investigated the deregulation of various biological pathways in a total of 47 peripheral blood mononuclear cell (PBMC) samples from patients with sporadic PD, patients heterozygous for the LRRK2 G2019S mutation compared to healthy controls. We found that the deregulation patterns were indeed similar in PBMCs obtained from patients with sporadic PD and from LRRK2 G2019S carriers, with dysfunctions in mitochondrial pathways, cell survival signaling, cancerization, endocytosis signaling and iron metabolism. Analysis of our PBMC data and other publicly available transcriptome datasets (for whole blood samples) showed that deregulation of the immune system, endocytosis and eukaryotic initiation factor 2 (EIF2) signaling are the main features of transcriptome profiles in PD (since they are also present in the transcriptome of dopaminergic neurons from patients). Transcriptome analysis of PBMCs is thus valuable for (i) characterizing the pathophysiological pathways shared by genetic and sporadic forms of PD and (ii) identifying potential biomarkers and therapeutic targets. This minimally invasive approach opens up tremendous perspectives for better diagnosis and therapy of neurodegenerative diseases because it can be applied from the earliest stages of the disease onwards.

## Introduction

Parkinson's disease (PD) is the most prevalent movement disorder in elderly adults. Although most cases of PD are sporadic (i.e. caused by complex interactions between many different genetic and environmental factors), the discovery of inherited forms of PD opened up opportunities for studying genetic aspects of the disease in detail. The first gene to be identified was the  $\alpha$ -synuclein gene (SNCA), which is affected by missense mutations and multiplications (Singleton *et al.*, 2013). Although these mutations are rare, this  $\alpha$ -synuclein has a key role in the pathogenesis of PD. Indeed, the spread of  $\alpha$ -synuclein and  $\alpha$ -synuclein aggregation are key events in PD. Autosomal dominant PD has also been linked to mutations in the leucine-rich repeat kinase 2 (*LRRK2*) gene, with the G2019S substitution being the most common (Singleton *et al.*, 2013). However, few studies have explored *LRRK2*'s role in peripheral tissues. Peripheral blood mononuclear cells (PBMCs) constitute an easily accessible window onto the multi-organ transcriptome (Kohane & Valtchinov, 2012). We have previously used the transcriptional profiling of PBMCs in proof-of-concept experiments designed to reveal perturbations in pathways known to be involved in PD-related neurodegeneration; this included the ubiquitin-proteasome system and the mitochondrial oxidation system (Mutez *et al.*, 2011). Since patients with sporadic PD and patients with *LRRK2*-linked PD display the classical clinical phenotype, we sought to determine the extent to which biological pathways were altered in these two populations by applying a hypothesis-free genetic approach. We thus described and compared gene expression profiles in subjects, who were heterozygous for the *LRRK2* G2019S mutation, patients with sporadic PD versus healthy control subjects. We identified similarities and differences in the transcriptome profiles of these PD subtypes and also compared our data with publically available transcriptome datasets for blood cells or dopaminergic neurons from PD subjects and controls. We hypothesized that better knowledge of deregulated genes in these two groups would shed further light on the pathophysiology of PD.

## Methods

### *Subjects*

Participants were recruited from the Movement Disorders Unit at the Lille University Medical Center (Lille, France). The study was approved by the investigational review board and all the subjects gave their informed consent to participation. All the subjects were examined by physicians trained in the diagnosis of PD (on the basis of the United Kingdom Parkinson's Disease Society Brain Bank criteria (Gibb *et al.*, 1988)). In the PD group, we studied 9 individuals who were heterozygous for the *LRRK2* G2019S mutation (mean age:  $66.2 \pm 17$ ; 4 males and 5 females) and 20 individuals with sporadic PD

(mean age:  $62.3 \pm 8$ ; 9 males and 11 females). Forty unrelated controls (20 males, 20 females, mean age  $52.9 \pm 19$ ) were recruited from volunteers free of any personal or familial neurological disorders and were sampled to prepare a pool of control RNA. This pool consisted of equimolar amount of RNA in order to better ascertain the variability of the control group at a lower cost and to assess the reproducibility of the microarrays. None of the controls had a personal or family history of neurological disorders. Furthermore, 19 of the 40 control participants (7 males and 12 females; mean age:  $59.5 \pm 14$ ) were studied individually. No pathogenic LRRK2 mutations were identified in the patients with sporadic PD and no LRRK2 G2019S mutations were identified in the control participants (Ross *et al.*, 2011).

#### *Microarray procedure*

Total RNA was extracted from the subjects' PBMCs. The microarray procedure (with one-color whole-human genome 44K arrays from Agilent, Santa Clara, CA, USA) has been described previously (Mutez *et al.*, 2011). We performed a new transcriptome analysis with the data used in the previous analysis ie the 9 PD patients with the G2019S *LRRK2* mutation and the 8 controls including the control pool (Mutez *et al.*, 2011), and we added 12 other control samples and 20 sporadic PD samples. Raw microarray data were deposited in the Gene Expression Omnibus (GEO) database, [www.ncbi.nlm.nih.gov/geo](http://www.ncbi.nlm.nih.gov/geo) (accession number GSE49126). Data analyses were performed with GeneSpring GX 11.0 software (Agilent, Santa Clara, CA, USA). A volcano plot was used to identify differentially expressed genes (with a  $\geq 1.2$  fold-change (FC) and  $p < 0.05$  as the threshold for statistical significance in Welch's *t*-test with Benjamini-Hochberg correction for multiple testing). Genes with significantly differential expression were submitted to Ingenuity Pathways Analysis (IPA, [www.ingenuity.com](http://www.ingenuity.com)) to identify deregulated pathways and toxicity lists, as described elsewhere (Mutez *et al.*, 2011). In order to validate the microarray data, we used an Illumina platform (the Human HT-12 v4 Expression BeadChip) to perform a second transcriptome analysis of RNA samples from a 55-year-old patient with PD (in triplicate) and 18 control participants. We used the same procedure to identify significantly deregulated genes ( $FC \geq 1.2$  and  $p < 0.05$  in Welch's *t*-test). Of the 367 genes found to be deregulated in both platforms, 79% display deregulation in the same direction (i.e. over or under-expression) as in the first analysis (data not shown). The 21% difference can be explained by the fact that Agilent and Illumina microarray platforms use different technologies (slides and beads, respectively) and probe lengths and thus may target different transcripts. However, the high overall degree concordance suggests that our results are valid.



### *Comparison with transcriptome datasets obtained from whole blood of PD subjects and controls*

We compared our data with two publicly available transcriptome datasets for PD, with a view to identifying deregulated genes as candidate biomarkers and/or with differential exon expression (Gene Expression Omnibus (GEO) datasets; [www.ncbi.nlm.nih.gov/gds](http://www.ncbi.nlm.nih.gov/gds)). The two datasets corresponded to (i) a whole-genome expression array (the Affymetrix Human Genome U133A Array) analysis of peripheral blood samples from 50 PD subjects and 22 healthy controls (GEO accession number GSE6613; Scherzer *et al.*, 2007) and (ii) an exon array analysis (the Affymetrix Human Exon 1.0 ST array) of peripheral blood samples from 17 patients with PD subjects and 11 healthy controls (GEO accession number GSE18838; Shehadeh *et al.*, 2010). We re-analyzed the data with the GEO2R web tool (<http://www.ncbi.nlm.nih.gov/geo/geo2r/>) using the Linear Models For Microarray Analysis (LIMMA) package in R software, in order to identify genes that were expressed differentially in each of the two datasets. The threshold for statistical significance was set to  $p < 0.05$ ; furthermore, only differentially expressed genes with a  $FC \geq 1.2$  were submitted to IPA.

## **Results**

*Deregulated pathways in PBMCs from patients with sporadic PD and LRRK2 mutation carriers: highlight on pathways known to be involved in PD physiopathology.*

We used GeneSpring GX 11.0 and IPA software to study the transcriptome profiles of PBMCs from patients with sporadic PD and from LRRK2 mutation carriers. The most significantly deregulated pathway in the group of patients with sporadic PD (Supplementary Table 1) was AMP-activated protein kinase (AMPK) signaling, which was recently linked to the presence of Lewy-body-bearing neurons and  $\alpha$ -synuclein accumulation in PD (Jiang *et al.*, 2013). Other notably deregulated pathways were involved in inflammation and immune processes (Glucocorticoid receptor signaling, airway inflammation in asthma signaling, Interleukin (IL)-4, IL-9, IL-15 signaling, the T cell receptor signaling and CCR3 signaling), cell survival (Myc-mediated apoptosis, induction of apoptosis by HIV, 14-3-3 signaling, HIF $\alpha$  signaling, PI3K/AKT and ERK/MAPK signaling), cancerization (in prostate cancer, breast cancer, lung cancer and chronic myeloid leukemia signaling), regulation of protein production (including eukaryotic initiation factor 2 (EIF2) signaling, DNA methylation and transcriptional repression signaling) and endocytosis (virus entry via endocytic pathways).

In the group of patients with LRRK2 mutation (Supplementary Table 2), we found peripheral deregulation of pathways involved in inflammation and immune processes (leukocyte extravasation signaling, cytotoxic T lymphocytes, CXCR4 signaling, CD28 signaling in T Helper cells, B cell

development and IL2 signaling), cell survival and the cell cycle (p53, Cdc42 signaling and PI3K/AKT signaling), cancer processes (molecular mechanisms of cancer, renal cell carcinoma, prostate cancer, endometrial cancer and chronic myeloid leukemia), the regulation of translation (EIF2 signaling and the regulation of eukaryotic translation initiation factor 4 (EIF4) and p70S6K signaling) and endocytosis (clathrin-mediated endocytosis signaling). It is noteworthy that some of the most severely deregulated pathways have been directly linked to the physiopathology of LRRK2 in various models; these include EIF2 signaling via the regulation of translation (Gehrke *et al.*, 2010), mTOR signaling via the regulation of autophagy (Ferree *et al.*, 2012), actin cytoskeleton signaling (Meixner *et al.*, 2011) and GTPase signaling via RhoA and Rac signaling (Taymans *et al.*, 2012).

Since the clinical and pathological phenotypes of sporadic PD subjects and *LRRK2* mutation carriers are almost indistinguishable (Healy *et al.*, 2008), we focused on the eight most common canonical pathways in these two groups of PD patients (Table 1). Our analysis highlighted the peripheral deregulation of pathways typically involved in the neurodegenerative process in PD: (i) mitochondrial dysfunction and oxidative phosphorylation, (ii) EIF2 signaling, (iii) Rac signaling (as in the *LRRK2* analysis mentioned above (Taymans *et al.*, 2012), (iv) deregulation of cancer signaling (chronic myeloid leukemia and prostate cancer), with an overlap with PD in terms of the regulation of cell survival in general and the key PI3K/AKT signaling pathway in particular (Devine *et al.*, 2011) and, lastly, (v) purine metabolism and its end product urate (which might constitute both a marker of the development and progression of PD and a potentially neuroprotective treatment (Chen *et al.*, 2012)). Thus, it was possible to link these eight common pathways deregulated in PBMCs to the PD process and the physiopathology of *LRRK2*.

#### *Common deregulated genes and pathways in the two PD groups: endocytosis and iron metabolism.*

In order to gain a more comprehensive view of the deregulated pathways, we identified the genes that were most significantly deregulated ( $FC \geq 1.2$ , Welch's *t*-test  $p < 0.05$  with Benjamini-Hochberg correction) in both patients with sporadic PD and *LRRK2* mutation carriers. Analysis of the 120 identified genes revealed that EIF2 signaling was the most severely deregulated canonical pathway ( $p = 0.00001$ ). We also observed over-representation of genes involved in iron metabolism and heme biosynthesis (such as *ALAS2*, *EPB42*, *GYPB*, *HBD*, *HBM* and *SLC4A1*; Supplementary Table 3). In PD, perturbation of genes involved in iron metabolism has already been described (for the substantia nigra) (Mandel *et al.*, 2006). Expression of *ALAS2* and *HBD* is controlled by the transcription factor GATA1, which also regulates expression of *SNCA* (Scherzer *et al.*, 2008). This suggests the existence of a link between GATA1 and PD. It is also noteworthy that *ALAS2* was down-regulated in both groups of

PD patients, confirming our previous RT-PCR-based observations in PBMCs obtained from *LRRK2* mutation heterozygous individuals (Mutez *et al.*, 2011).

*Deregulated genes and pathways in peripheral tissues in Parkinson's disease: the contributions of EIF2 signaling, virus entry via endocytic pathways and leukocyte extravasation signaling.*

Few studies have investigated the deregulation gene expression in peripheral tissues in PD (Cooper-Knock *et al.*, 2012). To identify common deregulated genes, we compared the genes found to be deregulated in our present study with those identified in two transcriptome studies of whole blood samples from PD patients (Scherzer *et al.*, 2007; Shehadeh *et al.*, 2010). Importantly, genes involved in iron metabolism such as *ALAS2*, *EPB42*, *SELENBP1* and *SLC4A1* were deregulated in all three studies (Supplementary Table 3). These four genes thus deserve further investigation, in order to assess their potential as peripheral biomarkers for PD. In addition to the deregulation of gene expression, we used IPA software to look for canonical pathways that were common to all three studies. We identified common deregulated canonical pathways in at least three of the four PD groups (*LRRK2* mutation carriers, subjects with sporadic PD from our study and subjects with sporadic PD from Scherzer *et al.*, 2007 and Shehadeh *et al.*, 2010). The three most commonly affected pathways were EIF2 signaling, virus entry via endocytic pathways and leukocyte extravasation signaling and might correspond to new pathophysiological mechanisms in PD.

*Common cell signaling alterations in PBMCs and dopaminergic neurons: the preeminent role of cell survival pathways.*

A recent literature review listed the gene expression profiling studies in human neurodegenerative diseases in general and PD in particular (Cooper-Knock *et al.*, 2012). Most of the studies focused on gene expression in the substantia nigra or other brain and were based on homogenized tissue (i.e. containing many different cell types, the involvement of which in the physiopathology PD may differ). More specifically, two groups have studied profiled gene expression in dopaminergic neurons obtained by laser capture microdissection (Simunovic *et al.*, 2010; Elstner *et al.*, 2011). We compared the latter two profiles with the profiles observed in our present study. Elstner *et al.* had used IPA software to find canonical pathways in their dataset and that produced by Simunovic *et al.* Of the 25 common deregulated canonical pathways in Elstner *et al.*'s and Simunovic *et al.*'s studies, 19 were also deregulated in PBMCs from our *LRRK2* mutation carriers or individuals with sporadic PD (Supplementary Tables 1 and 2). Pathways usually associated with PD (such as mitochondrial dysfunction and the protein ubiquitination pathway) were profoundly deregulated in the PBMCs in our present study.

## Discussion

Peripheral blood mononuclear cells are often used to identify cellular dysfunctions classically associated with the physiopathology of PD, such as a reduction in proteasome activity (Ullrich *et al.*, 2010), mitochondrial dysfunction (Shinde & Pasupathy, 2006; Zheng *et al.*, 2010), increased reactive oxygen species production (Prigione *et al.*, 2006), low dopamine content and low dopamine transporter immunoreactivity (Colombo *et al.*, 2003). In the present study, we investigated the deregulation of gene expression in PBMCs obtained from two groups of PD subjects; patients with sporadic PD and patients with a genetically determined form of PD (due to mutation of the *LRRK2* gene). In each group, we found that genes and pathways involved in the physiopathology of PD or mutant *LRRK2* were indeed deregulated. Our results emphasize the potential value of PBMC transcriptome profiling in studies of peripheral tissues in PD.

*The presence of immune and inflammatory disturbances emphasizes that PD is a systemic disorder.*

We observed altered immune and inflammatory processes in different subtypes of PD. These deregulations are prominent in the PD brain, with elevated number of HLA-DR-positive activated microglia (Imamura *et al.*, 2003). In this regard, the HLA-DR histocompatibility complex class II receptor locus and recently genes involved in the 'regulation of leucocyte/lymphocyte activity' and also 'cytokine-mediated signaling' have also been identified in genome-wide association studies of PD (Nalls *et al.*, 2011; Holmans *et al.*, 2013). Furthermore, immune system involvement in PD was suggested by deregulated expression/release of pro-inflammatory cytokines and infiltration of CD4+ T lymphocytes into the brain (Brochard *et al.*, 2009; Durrenberger *et al.*, 2012). An increase in numbers of natural killer cells and a role for  $\alpha$ -synuclein and LRRK2 in immune cell functions have also been reported recently (Niwa *et al.*, 2012; Thevenet *et al.*, 2011; Gardai *et al.*, 2013). Interestingly, leukocyte extravasation signaling was found to be deregulated in our present dataset and in the two whole-blood transcriptome datasets (Scherzer *et al.*, 2007; Shehadeh *et al.*, 2010). While it could be argued that this type of deregulation is due to the cell type in question, our data suggest that these mononuclear cells contribute to the disease process and should thus be investigated further in the context of PD.

*Endocytosis as a commonly deregulated pathway in PD.*

The neuro-inflammatory processes observed in PD (i.e. microglia activation and lymphocyte infiltration) are activated by the release of  $\alpha$ -synuclein by neurons (Alvarez-Erviti *et al.*, 2011). In fact,  $\alpha$ -synuclein can propagate from cell to cell via endocytic processes (Angot *et al.*, 2012). Here, we

observed that the canonical pathway for virus entry via endocytosis was deregulated in PBMCs from patients with PD. It has already been demonstrated that  $\alpha$ -synuclein and LRRK2 are involved in the modulation of endocytosis and, in particular, the trafficking of synaptic vesicles. Indeed,  $\alpha$ -synuclein promotes clathrin-mediated endocytosis and synaptic vesicle recycling (Ben Gedalya *et al.*, 2009) and LRRK2 controls synaptic endocytosis via its kinase activity (Matta *et al.*, 2012). Other genes involved in PD are also implicated in lysosomal vesicular trafficking (Dehay *et al.*, 2013), these include the ATPase type 13A2 (ATP13A2), identified in a Mendelian form of Parkinsonism and the glucosidase, beta, acid (GBA) major susceptibility gene (Dehay *et al.*, 2013).  $\alpha$ -synuclein aggregates have been shown to induce the rupture of lysosomes after endocytosis in neuronal cell lines or after cell-to-cell transfer (Freeman *et al.*, 2013). Modulation of endocytosis and vesicular trafficking may contribute to the spread of  $\alpha$ -synuclein and thus appears to be an important potential therapeutic target.

#### *EIF2 signaling: a potential therapeutic target in PD?*

To date, research on the disease mechanisms in PD has focused on neuronal death, inflammatory phenomena and, most recently,  $\alpha$ -synuclein spreading. Using a PBMC transcriptome approach, we observed alterations of genes and pathways associated with these three mechanisms. Our results also highlighted EIF2 signaling as a new candidate pathway associated with PD: this was the most severely deregulated pathway in PBMCs from patients with sporadic PD and LRRK2 mutation carriers and in whole blood samples and dopaminergic neuronal cells from other PD populations (Scherzer *et al.*, 2007; Shehadeh *et al.*, 2010; Simunovic *et al.*, 2010; Elstner *et al.*, 2011). In a post-mortem transcriptome analysis of the substantia nigra from individuals with incidental Lewy body disease, patients with PD and age-matched controls; EIF2 signaling and mTOR signaling were the most severely deregulated pathways. Indeed, the EIF2 pathway is already significantly altered in early-phase PD (Dijkstra *et al.*, 2012; Dijkstra *et al.*, submitted). Impaired EIF2 signaling has been linked to the physiopathology of LRRK2 because the latter regulates protein synthesis via interaction with the microRNA-pathway and eukaryotic translation initiation factor 2C (also known as Argonaute 2 protein) (Gehrke *et al.*, 2010). EIF2 is mainly involved in the initiation phase of translation and translation regulation has been linked to other forms of parkinsonism. Mutations in Eukaryotic Translation Initiation Factor 4-Gamma 1 (*EIF4G1*) were recently linked to PD (Chartier-Harlin *et al.*, 2011). Mutations in Protein Kinase Interferon-Inducible Double-Stranded RNA-Dependent Activator (PRKRA), a protein sensing cellular stress and virus infection, have provided a genetic pointer to the PKR/EIF2 pathway and translational regulation as a mechanism involved in the etiology of a complicated parkinsonian/dystonia disorder called DYT16 (Camargos *et al.*, 2008). EIF2 regulates protein synthesis in response to various types of stress and, in particular, the unfolded protein

response (UPR) in the endoplasmic reticulum. Mutation of *LRRK2* leads to dopaminergic neuronal death via endoplasmic reticulum stress (Yuan *et al.*, 2011). Activation of the UPR has been documented in the substantia nigra of PD individuals with phosphorylated protein kinase RNA-like endoplasmic reticulum kinase (PERK, a key mediator of the UPR pathway) and phosphorylated EIF2- $\alpha$  immunoreactivity (Hoozemans *et al.*, 2007). A combination of impaired autophagy and UPR activation led to the accumulation of  $\alpha$ -synuclein. Interestingly, a specific inhibitor of the kinase PERK, prevented UPR-mediated translational repression linked to EIF2- $\alpha$  phosphorylation in prion disease. Remarkably, the compound prevented neurodegeneration both at the pre-clinical and clinical stages of the disease, despite continuing accumulation of misfolded prion protein (Moreno *et al.*, 2013). Future research must pinpoint EIF2's involvement in PD and establish whether this factor might constitute a new therapeutic target.

*Common cell signaling alterations in PBMCs and dopaminergic neurons highlight cell survival pathways.*

Impairments in pathways involved in immune functions, inflammation, endocytosis, the control of energy metabolism and cell survival in response to different types of stress (mitochondrial dysfunction, oxidative phosphorylation, PI3K/AKT signaling, mTOR signaling, EIF4 and p70S6K signaling and EIF2 signaling) were observed in both dopaminergic neurons and PBMCs. It is noteworthy that immune alterations (including leukocyte extravasation signaling), endocytosis perturbations (with virus entry via endocytic pathways and clathrin-mediated endocytosis signaling) and EIF2 signaling were also altered in the substantia nigra at Braak  $\alpha$ -synuclein stages 1-2 (Dijkstra *et al.*, 2012; Dijkstra *et al.*, submitted). Considered as a whole, these overall changes are consistent with experiments in cell-based and animal models, in which  $\alpha$ -synuclein misfolding initiates a neurodegenerative cascade with prion-like propagation. This event can be triggered by genetic mutations and/or oxidative stress and leads to an inflammatory response, a UPR and downstream perturbations in translation related to EIF2. Since translation is an energy-intensive process, these perturbations will influence cell survival pathways (see Figure 1). The key molecular events in this cascade might also be therapeutic targets. Indeed, it has already been shown that inhibition of mTOR via the activation of EIF4E binding protein 4E-BP prevents dopaminergic neuron loss (Tain *et al.*, 2009).

*Overlap in peripheral and central gene expression when comparing genetically determined and sporadic PD.*

Our present results emphasize the value of studying peripheral tissues in neurodegenerative disorders because these tissues can also provide information on impaired processes. Peripheral cells can be acquired in a minimally invasive manner and at any stage of the disease; in contrast, dopaminergic neurons can only be obtained post-mortem in end-stage disease with confounding deregulations of terminal state. The deregulated biological pathways observed in PBMCs appear to reflect a number of key events in the PD mechanism because they were observed in both sporadic and genetically determined forms of PD. In the genetically determined forms, the processes may be exacerbated by the presence of a mutation in each of the organism's cells. However, a similarly deregulated transcriptome profile is observed in sporadic forms of PD for which interaction between genetic susceptibility and environmental factors are suspected to be involved. Future research studies must seek to differentiate between processes linked to neurodegeneration and those linked to genomic changes. We are now recruiting a larger number of (i) patients with sporadic PD, (ii) patients presenting mutations in autosomal dominant PD genes and (iii) patients with other neurodegenerative disorders (such as multiple systemic atrophy, progressive supranuclear palsy, Alzheimer's disease, and Huntington's disease), in order to better understand the overlap between deregulations in the peripheral tissues and in the central nervous system (CNS). A recent review of transcriptome data (Cooper-Knock *et al.*, 2012) found that similar functional categories of deregulated genes (including those involved in neuro-inflammation and intracellular signaling) were found in CNS tissues, PBMCs and peripheral leukocytes obtained from patients with neurodegenerative diseases. Furthermore, the inclusion of presymptomatic patients with autosomal dominant transmission of harmful mutations might also provide other information on disease mechanisms and help to identify potential biomarkers (Mutez *et al.*, 2011). Nevertheless, our data emphasize the need to further investigate the effects of PD-related mutations and polymorphisms on immune functions (such as the innate immune response) and not just on CNS function (Gardai *et al.*, 2013). By exploring the transcriptome in PBMCs, it might be possible to investigate peripheral dysfunctions early in the disease. Taken as a whole, these considerations validate the analysis of peripheral tissues in PD and other neurodegenerative disorders.

In summary, we demonstrated the value of PBMC transcriptome profiling in PD. We observed deregulation of pathways that have already been described in dopaminergic neurons and that may contribute to the disease process. Deregulations of endocytosis, EIF2 signaling and immune pathways were identified in PBMCs obtained from individuals with sporadic PD and individuals with genetically determined PD. These molecular events highlighted key alterations that are common to different types of PD and may be potential therapeutic targets.

## Acknowledgments

We thank Dr. Lydie Larvor, Vincent Mouroux, Dr. Emilie Le Rhun and Dr. Nawal Waucquier for their help in recruiting patients with PD and controls and for their technical support. We thank Celine Villenet and Sabine Quief for assistance with the transcriptome experiments. We wish to thank the authors of other transcriptome studies for making their data available through public databases. This work was funded by the INSERM, the French Ministry of Health's PHRC program (PARKFANORD 2005/1914 and CONVERGENCE 2008-A00219-42), Lille University Medical Center and the Institut de Recherche pour le Cancer de Lille.

## References

- Alvarez-Erviti L, Couch Y, Richardson J, Cooper JM, Wood MJ.  $\alpha$ -synuclein release by neurons activates the inflammatory response in a microglial cell line. *Neurosci Res.* 2011;69:337-42.
- Angot E, Steiner JA, Lema Tomé CM, Ekström P, Mattsson B, Björklund A, *et al.*  $\alpha$ -synuclein cell-to-cell transfer and seeding in grafted dopaminergic neurons in vivo. *PLoS One.* 2012;7:e39465.
- Ben Gedalya T, Loeb V, Israeli E, Altschuler Y, Selkoe DJ, Sharon R.  $\alpha$ -synuclein and polyunsaturated fatty acids promote clathrin-mediated endocytosis and synaptic vesicle recycling. *Traffic.* 2009;10:218-34.
- Brochard V, Combadière B, Prigent A, Laouar Y, Perrin A, Beray-Berthat V, *et al.* Infiltration of CD4+ lymphocytes into the brain contributes to neurodegeneration in a mouse model of Parkinson disease. *J Clin Invest.* 2009;119:182-92.
- Camargos S, Scholz S, Simón-Sánchez J, Paisán-Ruiz C, Lewis P, Hernandez D, *et al.* DYT16, a novel young-onset dystonia-parkinsonism disorder: identification of a segregating mutation in the stress-response protein PRKRA. *Lancet Neurol.* 2008;7:207-15.
- Chartier-Harlin MC, Dachsel JC, Vilariño-Güell C, Lincoln SJ, Leprêtre F, Hulihan MM, *et al.* Translation initiator EIF4G1 mutations in familial Parkinson disease. *Am J Hum Genet.* 2011;89:398-406.
- Chen X, Wu G, Schwarzschild MA. Urate in Parkinson's disease: more than a biomarker? *Curr Neurol Neurosci Rep.* 2012;12:367-75.



- Chinta SJ, Mallajosyula JK, Rane A, Andersen JK. Mitochondrial  $\alpha$ -synuclein accumulation impairs complex I function in dopaminergic neurons and results in increased mitophagy in vivo. *Neurosci Lett*. 2010;486:235-9.
- Colombo C, Cosentino M, Marino F, Rasini E, Ossola M, Blandini F, *et al*. Dopaminergic modulation of apoptosis in human peripheral blood mononuclear cells: possible relevance for Parkinson's disease. *Ann N Y Acad Sci*. 2003;1010:679-82.
- Cooper-Knock J, Kirby J, Ferraiuolo L, Heath PR, Rattray M, Shaw PJ. Gene expression profiling in human neurodegenerative disease. *Nat Rev Neurol*. 2012;8:518-30.
- Dehay B, Martinez-Vicente M, Caldwell GA, Caldwell KA, Yue Z, Cookson MR, Klein C, *et al*. Lysosomal impairment in Parkinson's disease. *Mov Disord*. 2013;28:725-32.
- Devine MJ, Plun-Favreau H, Wood NW. Parkinson's disease and cancer: two wars, one front. *Nat Rev Cancer*. 2011;11:812-23.
- Dijkstra AA, Ingrassia A, Pattij T, Rozemuller AJM, Heutink P, Van de Berg WDJ. Transcriptome profiling implicates mTOR and EIF2 signaling in the pathogenesis of Parkinson's disease. Poster presentation, Society for Neuroscience, New Orleans, 2012.
- Dijkstra AA, Ingrassia A, Menezes R, van Kesteren R, Rozemuller AJM, Heutink P, van de Berg WDJ. Evidence for immune response, axonal dysfunction and reduced endocytosis preceding Lewy body pathology in the substantia nigra in Parkinson's disease. submitted.
- Durrenberger PF, Grünblatt E, Fernando FS, Monoranu CM, Evans J, Riederer P, *et al*. Inflammatory Pathways in Parkinson's Disease; A BNE Microarray Study. *Parkinsons Dis*. 2012;2012:214714.
- Elstner M, Morris CM, Heim K, Bender A, Mehta D, Jaros E, *et al*. Expression analysis of dopaminergic neurons in Parkinson's disease and aging links transcriptional dysregulation of energy metabolism to cell death. *Acta Neuropathol*. 2011;122:75-86.
- Ferree A, Guillily M, Li H, Smith K, Takashima A, Squillace R, *et al*. Regulation of physiologic actions of LRRK2: focus on autophagy. *Neurodegener Dis*. 2012;10:238-41.
- Freeman D, Cedillos R, Choyke S, Lukic Z, McGuire K, *et al*.  $\alpha$ -synuclein induces lysosomal rupture and cathepsin dependent reactive oxygen species following endocytosis. *PLoS One*. 2013;8:e62143.

- Freundt EC, Maynard N, Clancy EK, Roy S, Bousset L, Sourigues Y, Covert M, Melki R, Kirkegaard K, Brahic M. Neuron-to-neuron transmission of  $\alpha$ -synuclein fibrils through axonal transport. *Ann Neurol*. 2012;72:517-24.
- Gardai SJ, Mao W, Schüle B, Babcock M, Schoebel S, Lorenzana C, *et al*. Elevated  $\alpha$ -synuclein impairs innate immune cell function and provides a potential peripheral biomarker for Parkinson's disease. *PLoS One*. 2013;8:e71634.
- Gehrke S, Imai Y, Sokol N, Lu B. Pathogenic LRRK2 negatively regulates microRNA-mediated translational repression. *Nature*. 2010;466:637-41.
- Gibb WR, Lees AJ. The relevance of the Lewy body to the pathogenesis of idiopathic Parkinson's disease. *J Neurol Neurosurg Psychiatry*. 1988;51:745-52.
- Healy DG, Falchi M, O'Sullivan SS, Bonifati V, Durr A, Bressman S, *et al*. Phenotype, genotype, and worldwide genetic penetrance of LRRK2-associated Parkinson's disease: a case-control study. *Lancet Neurol*. 2008;7:583-90.
- Holmans P, Moskvin V, Jones L, Sharma M, International Parkinson's Disease Genomics Consortium, Vedernikov A, *et al*. A pathway-based analysis provides additional support for an immune-related genetic susceptibility to Parkinson's disease. *Hum Mol Genet*. 2013;22:1039-49.
- Hoozemans JJ, van Haastert ES, Eikelenboom P, de Vos RA, Rozemuller JM, Scheper W. Activation of the unfolded protein response in Parkinson's disease. *Biochem Biophys Res Commun*. 2007;354:707-11.
- Imamura K, Hishikawa N, Sawada M, Nagatsu T, Yoshida M, Hashizume Y. Distribution of major histocompatibility complex class II-positive microglia and cytokine profile of Parkinson's disease brains. *Acta Neuropathol*. 2003;106:518-26.
- Jiang P, Gan M, Ebrahim AS, Castanedes-Casey M, Dickson DW, Yen SH. Adenosine monophosphate-activated protein kinase overactivation leads to accumulation of  $\alpha$ -synuclein oligomers and decrease of neurites. *Neurobiol Aging*. 2013;34:1504-15.
- Kohane IS, Valtchinov VI. Quantifying the white blood cell transcriptome as an accessible window to the multiorgan transcriptome. *Bioinformatics*. 2012;28:538-545.

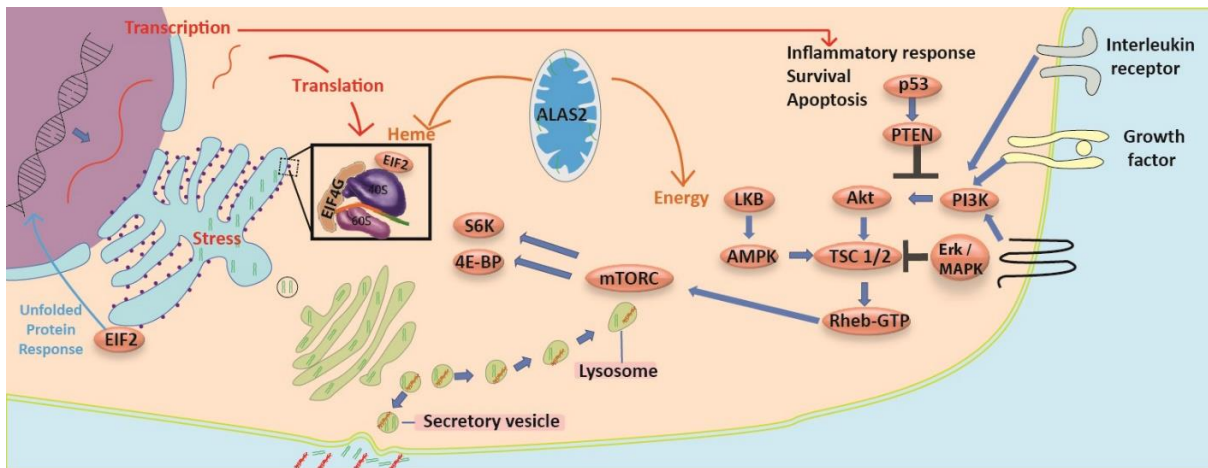
- Mandel S, Grunblatt E, Riederer P, Amariglio N, Jacob-Hirsch J, Rechavi G, *et al.* Gene expression profiling of sporadic Parkinson's disease substantia nigra pars compacta reveals impairment of ubiquitin-proteasome subunits, SKP1A, aldehyde dehydrogenase, and chaperone HSC-70. *Ann N Y Acad Sci.* 2005;1053:356-75.
- Matta S, Van Kolen K, da Cunha R, van den Bogaart G, Mandemakers W, Miskiewicz K, *et al.* LRRK2 controls an EndoA phosphorylation cycle in synaptic endocytosis. *Neuron.* 2012;75:1008-21.
- Meixner A, Boldt K, Van Troys M, Askenazi M, Gloeckner CJ, Bauer M, *et al.* A QUICK screen for Lrrk2 interaction partners-leucine-rich repeat kinase 2 is involved in actin cytoskeleton dynamics. *Mol Cell Proteomics.* 2011;10:M110.001172.
- Moreno JA, Halliday M, Molloy C, Radford H, Verity N, Axten JM, *et al.* Oral treatment targeting the unfolded protein response prevents neurodegeneration and clinical disease in prion-infected mice. *Sci Transl Med.* 2013;5:206ra138.
- Mutez E, Larvor L, Leprêtre F, Mouroux V, Hamalek D, Kerckaert JP, *et al.* Transcriptional profile of Parkinson blood mononuclear cells with LRRK2 mutation. *Neurobiol Aging.* 2011;32:1839-48.
- Nalls MA, Plagnol V, Hernandez DG, Sharma M, Sheerin UM, *et al.* International Parkinson Disease Genomics Consortium, Imputation of sequence variants for identification of genetic risks for Parkinson's disease: a meta-analysis of genome-wide association studies. *Lancet.* 2011;377:641-9.
- Niwa F, Kuriyama N, Nakagawa M, Imanishi J. Effects of peripheral lymphocyte subpopulations and the clinical correlation with Parkinson's disease. *Geriatr Gerontol Int.* 2012;12:102-7.
- Pan-Montojo F, Schwarz M, Winkler C, Arnhold M, O'Sullivan GA, Pal A, *et al.* Environmental toxins trigger PD-like progression via increased  $\alpha$ -synuclein release from enteric neurons in mice *Sci Rep.* 2012;2:898.
- Prigione A, Begni B, Galbussera A, Beretta S, Brighina L, Garofalo R, *et al.* Oxidative stress in peripheral blood mononuclear cells from patients with Parkinson's disease: negative correlation with levodopa dosage. *Neurobiol Dis.* 2006;23:36-43.

- Ross OA, Soto-Ortolaza AI, Heckman MG, Aasly JO, Abahuni N, Annesi G, *et al.* Association of LRRK2 exonic variants with susceptibility to Parkinson's disease: a case-control study. *Lancet Neurol.* 2011;10:898-908.
- Scherzer CR, Eklund AC, Morse LJ, Liao Z, Locascio JJ, Fefer D, *et al.* Molecular markers of early Parkinson's disease based on gene expression in blood. *Proc Natl Acad Sci U S A.* 2007;104:955-60.
- Scherzer CR, Grass JA, Liao Z, Pepivani I, Zheng B, Eklund AC, *et al.* GATA transcription factors directly regulate the Parkinson's disease-linked gene  $\alpha$ -synuclein. *Proc Natl Acad Sci USA.* 2008;105:10907-12.
- Shehadeh LA, Yu K, Wang L, Guevara A, Singer C, Vance J, *et al.* SRRM2, a potential blood biomarker revealing high alternative splicing in Parkinson's disease. *PLoS One.* 2010;5:e9104.
- Shinde S, Pasupathy K. Respiratory-chain enzyme activities in isolated mitochondria of lymphocytes from patients with Parkinson's disease: preliminary study. *Neurol India.* 2006 Dec;54:390-3.
- Simunovic F, Yi M, Wang Y, Macey L, Brown LT, Krichevsky AM, *et al.* Gene expression profiling of substantia nigra dopamine neurons: further insights into Parkinson's disease pathology. *Brain.* 2009;132:1795-809.
- Singleton AB, Farrer MJ, Bonifati V. The genetics of Parkinson's disease: progress and therapeutic implications. *Mov Disord.* 2013;28:14-23.
- Tain LS, Mortiboys H, Tao RN, Ziviani E, Bandmann O, Whitworth AJ. Rapamycin activation of 4E-BP prevents parkinsonian dopaminergic neuron loss. *Nat Neurosci.* 2009;12:1129-35.
- Taymans JM. The GTPase function of LRRK2. *Biochem Soc Trans.* 2012;40:1063-9.
- Thévenet J, Pescini Gobert R, Hooft van Huijsduijnen R, Wiessner C, Sagot YJ. Regulation of LRRK2 expression points to a functional role in human monocyte maturation. *PLoS One.* 2011;6:e21519.
- Ullrich C, Mlekusch R, Kuschnig A, Marksteiner J, Humpel C. Ubiquitin enzymes, ubiquitin and proteasome activity in blood mononuclear cells of MCI, Alzheimer and Parkinson patients. *Curr Alzheimer Res.* 2010;7:549-55.

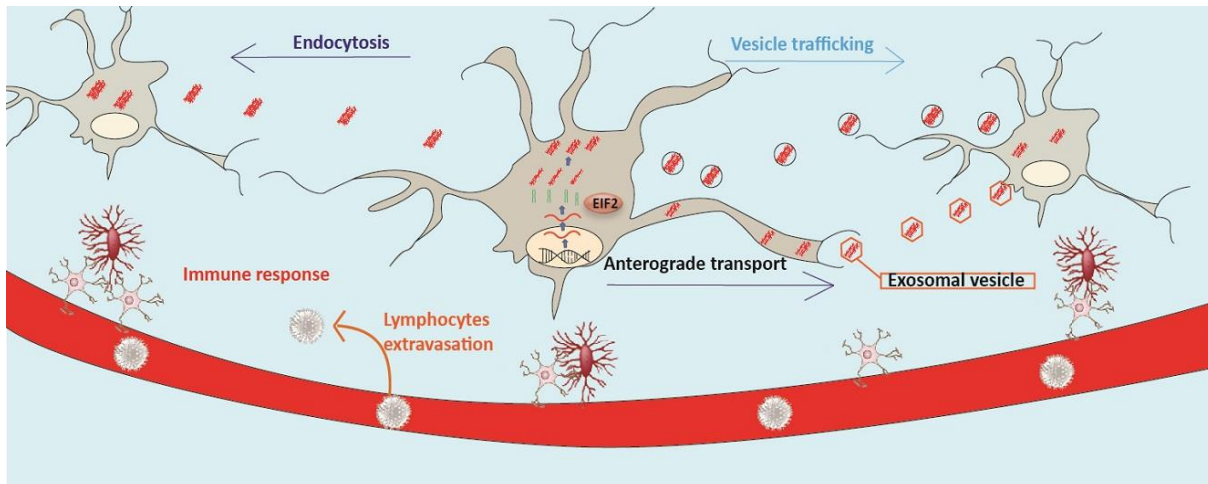
Yuan Y, Cao P, Smith MA, Kramp K, Huang Y, Hisamoto N, *et al.* Dysregulated LRRK2 signaling in response to endoplasmic reticulum stress leads to dopaminergic neuron degeneration in *C. Elegans*. *Plos One*. 2011;6:e22354.

Zheng B, Liao Z, Locascio JJ, Lesniak KA, Roderick SS, Watt ML, *et al.* PGC-1 $\alpha$ , a potential therapeutic target for early intervention in Parkinson's disease. *Sci Transl Med*. 2010;2:52ra73.

1.1



1.2



**Figure 1. Schematic overview of three mostly severely deregulated processes in PBMCs obtained from patients with sporadic PD and genetically determined PD (1) and their hypothetical involvement in the brain (2).**

(1) Deregulation of mitochondrial function (including alterations in oxidative phosphorylation, complex I and genes involved in heme biosynthesis), endocytic trafficking, immune & inflammatory processes and the main cell survival signaling pathways (ERK/MAPK and PI3K/AKT/mTOR/EIF2). The observation of impaired EIF2 signaling suggests that translation is perturbed in response to stress. In PBMCs from patients with PD, impaired leukocyte extravasation signaling may contribute to an inflammatory response by microglial cells.  $\alpha$ -synuclein at the inner mitochondrial membrane might also lead to mitochondrial fragmentation (Chinta *et al.*, 2010) and a decrease in the cell's energy levels. It is believed that this type of stress causes  $\alpha$ -synuclein to misfold within neurons.  $\alpha$ -synuclein

has been shown to delay endoplasmic reticulum (ER)-to-Golgi transport in cells (by antagonizing SNARE), disrupt lysosomes and thus contribute to inflammasome activation (Freeman *et al.*, 2013).

(2) Misfolded  $\alpha$ -synuclein is able to seed and recruit other  $\alpha$ -synuclein molecules. The progression of  $\alpha$ -synuclein fibrils might also be caused by neuron-to-neuron spreading of  $\alpha$ -synuclein aggregates following anterograde or retrograde axonal transport (Freundt *et al.*, 2012; Pan-Montojo *et al.*, 2012). The misfolded  $\alpha$ -synuclein can be released and transferred to the surrounding neurons by pore formation, exocytosis or exosome formation.

<b>Common Canonical Pathways</b>	<b>Sporadic PD patients</b>	<b>LRRK2 carriers</b>
Chronic Myeloid Leukemia Signaling	0.028	0.038
EIF2 Signaling	0.0025	5.01E-17
Mitochondrial Dysfunction	0.044	0.000032
Oxidative Phosphorylation	0.033	0.000038
PI3K/AKT Signaling	0.047	0.012
Prostate Cancer Signaling	0.0041	0.014
Purine Metabolism	0.017	0.00047
Rac Signaling	0.021	0.037

**Table 1.** Significant common canonical pathways in PBMCs of sporadic PD and LRRK2 G2019S PD patients.

The *p*-value (Fisher’s exact test) was indicated for each pathway.

**Supplementary Table 1.** Significant canonical pathways in PBMCs obtained from patients with sporadic PD. Pathways also found to be deregulated in a dopaminergic neuron transcriptome study (Eltner *et al.*, 2011) are shown in bold type. The *p*-value (in Fisher’s exact test) is indicated for each pathway.

**Supplementary Table 2.** Significant canonical pathways in PBMCs obtained from patients with PD induced by the LRRK2 G2019S mutation.

Pathways also found to be deregulated in a dopaminergic neuron transcriptome study (Eltner *et al.*, 2011) are shown in bold type. The *p*-value (in Fisher’s exact test) is indicated for each pathway.

**Supplementary Table 3.** Genes found to be deregulated in both patients with sporadic PD and heterozygous LRRK2 G2019S mutation carriers.

**Supplementary Table 1.**

Canonical pathways	<i>p</i> -value
<b>AMPK Signaling</b>	<b>0.0012</b>
<b>DNA Methylation and Transcriptional Repression Signaling</b>	<b>0.0018</b>
<b>Glucocorticoid Receptor Signaling</b>	<b>0.0024</b>
<b>EIF2 Signaling</b>	<b>0.0025</b>
Prostate Cancer Signaling	0.0041
<b>Breast Cancer Regulation by Stathmin1</b>	<b>0.0081</b>
Non-Small Cell Lung Cancer Signaling	0.01
<b>Ceramide Signaling</b>	<b>0.01</b>
Airway Inflammation in Asthma	0.012
N-Glycan Biosynthesis	0.013
β-alanine Metabolism	0.013
Urea Cycle and Metabolism of Amino Groups	0.015
Small Cell Lung Cancer Signaling	0.016
<b>Purine Metabolism</b>	<b>0.017</b>
PI3K Signaling in B Lymphocytes	0.018
Myc Mediated Apoptosis Signaling	0.019
<b>Glycolysis/Gluconeogenesis</b>	<b>0.019</b>
IL-9 Signaling	0.019
<b>Induction of Apoptosis by HIV1</b>	<b>0.020</b>
<b>14-3-3-mediated Signaling</b>	<b>0.021</b>
<b>Rac Signaling</b>	<b>0.021</b>
IL-15 Signaling	0.022
<b>Virus Entry via Endocytic Pathways</b>	<b>0.023</b>
Chronic Myeloid Leukemia Signaling	0.028
Telomere Extension by Telomerase	0.029
<b>Androgen Signaling</b>	<b>0.033</b>



Docosahexaenoic Acid (DHA) Signaling	0.033
<b>Oxidative Phosphorylation</b>	<b>0.033</b>
Erythropoietin Signaling	0.035
T Cell Receptor Signaling	0.036
Growth Hormone Signaling	0.038
IL-4 Signaling	0.038
CCR3 Signaling in Eosinophils	0.038
Nur77 Signaling in T Lymphocytes	0.04
Dopamine-DARPP32 Feedback in cAMP Signaling	0.041
Pyruvate Metabolism	0.042
<b>Mitochondrial Dysfunction</b>	<b>0.044</b>
HIF1 $\alpha$ Signaling	0.046
<b>ERK/MAPK Signaling</b>	<b>0.047</b>
<b>PI3K/AKT Signaling</b>	<b>0.047</b>

Supplementary Table 2.

Canonical pathways	<i>p</i> -value
<b>EIF2 Signaling</b>	<b>5.01E-17</b>
<b>mTOR Signaling</b>	<b>9.12E-09</b>
<b>Regulation of eIF4 and p70S6K Signaling</b>	<b>7.76E-07</b>
<b>Mitochondrial Dysfunction</b>	<b>3.23E-05</b>
<b>Oxidative Phosphorylation</b>	<b>3.89E-05</b>
<b>Ubiquinone Biosynthesis</b>	<b>0.00046</b>
<b>Purine Metabolism</b>	<b>0.00047</b>
<b>Protein Ubiquitination Pathway</b>	<b>0.001</b>
Regulation of Actin-based Motility by Rho	0.001
Antigen Presentation Pathway	0.0011
<b>Actin Cytoskeleton Signaling</b>	<b>0.0022</b>
<b>Estrogen Receptor Signaling</b>	<b>0.0022</b>
Integrin Signaling	0.0028
<b>RhoA Signaling</b>	<b>0.0028</b>
<b>Huntington's Disease Signaling</b>	<b>0.0044</b>

Neuregulin Signaling	0.0045
ErbB2-ErbB3 Signaling	0.0069
p53 Signaling	0.0075
RhoGDI Signaling	0.0085
<b>ILK Signaling</b>	<b>0.0087</b>
NGF Signaling	0.0087
Inositol Phosphate Metabolism	0.0093
<b>Germ Cell-Sertoli Cell Junction Signaling</b>	<b>0.01</b>
<b>Wnt/<math>\beta</math>-catenin Signaling</b>	<b>0.011</b>
IGF-1 Signaling	0.011
<b>PI3K/AKT Signaling</b>	<b>0.012</b>
<b>Renal Cell Carcinoma Signaling</b>	<b>0.012</b>
Telomerase Signaling	0.012
Prostate Cancer Signaling	0.014
<b>Cholecystokinin/Gastrin-mediated Signaling</b>	<b>0.015</b>
CTLA4 Signaling in Cytotoxic T Lymphocytes	0.016
Graft-versus-Host Disease Signaling	0.017
Phospholipase C Signaling	0.019
Autoimmune Thyroid Disease Signaling	0.019
Leukocyte Extravasation Signaling	0.021
iCOS-iCOSL Signaling in T Helper Cells	0.022
Molecular Mechanisms of Cancer	0.024
Neuroprotective Role of THOP1 in Alzheimer's Disease	0.024
<b>Cdc42 Signaling</b>	<b>0.026</b>
B Cell Development	0.026
RAR Activation	0.026
<b>CXCR4 Signaling</b>	<b>0.029</b>
Signaling by Rho Family GTPases	0.030
Endometrial Cancer Signaling	0.033
Oncostatin M Signaling	0.034
<b>p70S6K Signaling</b>	<b>0.035</b>
FLT3 Signaling in Hematopoietic Progenitor Cells	0.036

Production of Reactive Oxygen Species in Macrophages	0.036
<b>Rac Signaling</b>	<b>0.037</b>
IL-2 Signaling	0.038
Chronic Myeloid Leukemia Signaling	0.038
Glycerolipid Metabolism	0.039
Insulin Receptor Signaling	0.041
PPAR $\alpha$ /RXR $\alpha$ Activation	0.042
CD28 Signaling in T Helper Cells	0.043
Pantothenate and CoA Biosynthesis	0.045
Gap Junction Signaling	0.047
Pyrimidine Metabolism	0.047
<b>Clathrin-mediated Endocytosis Signaling</b>	<b>0.048</b>

**Supplementary Table 3.**

Symbol	Entrez Gene Name	LRRK2 mutation carriers		Sporadic PD subjects	
		Fold-change	Agilent Probe	Fold-change	Agilent Probe
<b>ALAS2</b>	aminolevulinate, delta-, synthase 2	-7.19	A_32_P22654	-8.13	A_32_P22654
		-6.75	A_32_P385587	-7.35	A_32_P385587
<b>EPB42</b>	erythrocyte membrane protein band 4.2	-4.71	A_23_P140675	-5.98	A_23_P140675
<b>GYPB</b>	Glycophorin B (MNS blood group)	-2.59	A_23_P212854	-2.58	A_23_P212854
		-1.19	A_24_P393644	-1.86	A_24_P393644
<b>HBD</b>	hemoglobin, delta	-2.77	A_23_P87346	-4.14	A_23_P87346
		-3.34	A_24_P75190	-4.90	A_24_P75190
<b>HBM</b>	hemoglobin, mu	-3.45	A_23_P15055	-3.92	A_23_P15055
<b>SELENBP1</b>	selenium binding protein 1	-3.02	A_23_P74619	-3.08	A_23_P74619
		-6.25	A_23_P86021	-5.12	A_23_P86021
<b>SLC4A1</b>	solute carrier family 4, anion exchanger, member 1 (erythrocyte membrane protein band 3)	-1.72	A_23_P77980	-2.13	A_23_P77980
		-6.80	A_23_P89380	-1.30	A_23_P89380
		-2.75	A_24_P385190	-5.91	A_24_P385190

### 3.2. **Chapitre 2** : Étude du transcriptome de cellules mononucléées périphériques sanguines de sujets porteurs de mutations du gène *ATXN2*

L'augmentation du nombre de triplets CAG au sein du gène *ATXN2* a longtemps été définie comme responsable de l'ataxie spinocérébelleuse de type 2 (patients SCA2c). Cependant elles sont également liées à la MP (patients SCA2p) et responsables d'un risque augmenté de développer la sclérose latérale amyotrophique (SLA), une maladie connue pour être associée à des perturbations de l'épissage. Physiologiquement compris entre 13 et 31, le nombre de répétitions oscille entre 27 et 33 chez les patients SCA2p et ceux atteints de SLA et augmente jusqu'à atteindre au maximum 200 répétitions chez les patients SCA2c. Alors que leur taille et leur structure ont été étudiées et différent entre les SCA2c et les SCA2p, le mécanisme par lequel ces mutations conduisent à des phénotypes différents n'a pas encore été élucidé.

*ATXN2* code la protéine ataxine-2, principalement connue pour son rôle dans le métabolisme des ARNs. L'une des premières fonctions qui lui a été attribuée est un rôle dans l'épissage des ARNs du fait de son interaction avec les facteurs d'épissage RBM9, RBPMS, TDP-43 et FUS et parce qu'elle possède en N-terminal un domaine Lsm/Lsm-AD caractéristique de ces facteurs. Elle possède également un domaine PAM2 en C-terminal qui lui permet d'interagir avec la protéine PABPC1 (polyA Binding Protein Cytoplasmic 1), une protéine impliquée dans la dégradation des ARNs par le nonsense-mediated decay (NMD), les granules de stress et la traduction. D'autres fonctions notamment dans le métabolisme calcique et lipidique lui ont été attribuées mais ont été rarement évoquées. Pour tenter de comprendre pourquoi les sujets SCA2c et SCA2p présentent des phénotypes différents, nous avons émis l'hypothèse que les différences au niveau des répétitions de triplets pouvaient soit altérer la même fonction de l'ataxine-2 mais dans des sens différents (activation d'un côté, inhibition de l'autre par exemple) ou au contraire agir au niveau de fonctions différentes de la protéine. Etant donné le rôle de l'ataxine-2 dans le métabolisme des ARNs et notre intérêt pour l'épissage, nos résultats sont orientés vers l'analyse de ce métabolisme, ce qui n'exclue pas la possibilité que d'autres fonctions de l'ataxine-2 méritent notre attention.

Ainsi, afin d'apporter des éléments de réponse concernant les mécanismes perturbés en réponse à des mutations de l'*ATXN2*, nous avons effectué une analyse du transcriptome de cellules mononucléées périphériques sanguines de 7 sujets SCA2p, 6 sujets SCA2c, 7 parkinsoniens sporadiques sans mutation et 13 sujets témoins. Par l'utilisation de puces d'expression, nous avons observé les profils d'expression de chaque groupe de sujets et identifié les variations d'expression de gènes entre les malades et les témoins. De façon similaire à l'analyse présentée dans le chapitre 1, l'étude des voies canoniques a été effectuée à l'aide du logiciel IPA®. En complément de cet outil,

nous avons utilisé les logiciels GSEA (<http://www.broadinstitute.org/gsea/index.jsp>) et Toppgene (<https://toppgene.cchmc.org/>) qui nous ont respectivement permis d'identifier les processus biologiques dans lesquels ces gènes étaient impliqués et de déterminer dans quel aspect du métabolisme des ARNs ces gènes intervenaient. Récemment Yokoshi *et al.* ont montré que l'ataxine-2 interagissait avec la région 3' non traduite des ARNm pour réguler leur stabilité et ont cherché à savoir si ce rôle était lié à son interaction avec PABPC1. Ainsi, en éliminant ou non le domaine PAM2 de l'ataxine-2, ils ont observé que certains ARNm se fixaient à l'ataxine-2 de manière spécifiquement dépendante de son interaction avec PABPC1 alors que d'autres ne se fixaient qu'en absence de la liaison avec ce partenaire. Nous avons donc soumis à Webgestat (<http://bioinfo.vanderbilt.edu/webgestalt/>) les listes de gènes appartenant aux 2 catégories de cibles de l'ataxine-2 identifiées par Yokoshi *et al.* (dépendantes ou non de la liaison avec PABPC1) pour identifier les voies KEGG (interactions moléculaires) associées. Ces voies ont ensuite été comparées à nos données d'analyse de voie IPA afin d'évaluer si les variations d'expression génique observées sur nos puces d'expression pouvaient être dépendante de PABPC1. Enfin, nous avons recherché la variation d'expression de gènes d'intérêt identifiés d'après nos données, dans des données de puces d'expression obtenues à partir d'échantillons de *SNpc* de parkinsoniens versus des témoins.

Notre analyse des voies et processus biologiques liés aux gènes différentiellement exprimés entre les malades et les témoins a mis en évidence aussi bien des modifications spécifiques à chaque groupe que des dérèglements communs. Elle a mis en évidence des voies communes entre les parkinsoniens sporadiques et les patients SCA2p telles que la voie de signalisation de l'infiltration leucocytaire, les voies de survie cellulaire, l'inflammation et du cytosquelette, en accord avec leurs similitudes phénotypiques. Les variations d'expression des gènes *ALAS2*, *HBD*, *HDM* et *SELENBP1* déjà identifiées chez les patients parkinsoniens au cours d'une de nos précédentes études sont également retrouvées chez les patients parkinsoniens sporadiques et les patients SCA2p de cette analyse. De façon intéressante, nous avons observé des changements d'expression de gènes de la voie de signalisation de la SLA chez l'ensemble des parkinsoniens mais pas chez les sujets SCA2c. Connaissant l'importance des perturbations du métabolisme des ARNs dans la SLA, nous avons supposé que l'identification de cette voie pouvait évoquer un lien entre le métabolisme des ARNs et la MP. Cette modification de la voie de la SLA est également présente dans l'analyse des voies KEGG des gènes ciblés par l'ataxine-2, uniquement lorsque celle-ci peut interagir avec PABPC1. PABPC1 permettant de discriminer entre les ARNm cibles de l'ataxine-2, il se pourrait que des mécanismes dépendants de cette protéine permettent de séparer les parkinsoniens (SCA2p et sporadiques) des SCA2c.

Étant donné le rôle de PABCP1 dans le métabolisme des ARNs, nous avons recherché si des perturbations de ce métabolisme étaient présentes chez les différents groupes de malades. L'analyse par GSEA des processus biologiques impactés par les variations d'expression génique a confirmé l'existence de dérèglements du métabolisme des ARNs chez les patients SCA2c, les patients SCA2p et les parkinsoniens sporadiques, soutenant l'hypothèse que les différentes mutations de l'ataxine-2 jouent sur son rôle dans le métabolisme des ARNs mais de façon différente. L'analyse détaillée des différents aspects du métabolisme des ARNs effectuée via Toppgene a permis de révéler que les processus du métabolisme des ARNs perturbés dans les différents groupes de malades impliquaient préférentiellement des aspects en lien avec la queue polyA des ARNm aussi bien chez les sujets SCA2c, les sujets SCA2p et les sporadiques. De façon intéressante, la comparaison des gènes du métabolisme des ARNs dérégulés dans chaque groupe, a montré une plus grande similitude entre les gènes dérégulés chez les parkinsoniens SCA2p et les cas sporadiques par rapport aux autres comparaisons (SCA2c *versus* sporadiques et SCA2c *versus* SCA2p). Ces résultats suggèrent que bien que le métabolisme des ARNs soit perturbé chez l'ensemble des malades les gènes impliqués dans ces perturbations seraient spécifiques du phénotype.

Nous avons donc sélectionné des gènes liés au métabolisme des ARNs dont l'expression varie uniquement chez les parkinsoniens (SCA2p et sporadiques) ou uniquement chez les SCA2c afin de tester si les modifications du métabolisme des ARNs impliquant ces gènes, étaient spécifiques d'un phénotype donné. La sous-expression des gènes *hnRNPA3*, *hnRNPUL1* et *MBNL1* observée uniquement dans les données de puces de sujets SCA2c a été confirmée par RT-qPCR avec une diminution significative ( $p < 0,05$ ) de 35%, 21% et 40% respectivement. La sous-expression de *MBNL1* a également été observée dans les données de RT-qPCR de sporadiques *versus* témoins bien qu'elle n'ait pas été observée sur les données de puces correspondantes à ces sujets. Le gène *PABPC4*, un paralogue de *PABCP1*, est quant à lui significativement dérégulé chez les parkinsoniens sporadiques à la fois sur les données de puces et dans les résultats de RT-qPCR. En montrant des variations quantitatives spécifiques des sujets SCA2p et des SCA2c, ces résultats étayent l'hypothèse que les mutations de l'*ATXN2* induisent des modifications des mêmes processus mais de façon différente et le fait que les gènes dérégulés chez les malades SCA2p se retrouvent dérégulés de la même manière chez les cas sporadiques suggèrent que ces perturbations puissent être liées à l'émergence d'un phénotype parkinsonien. Du fait de son rôle dans la MP, nous avons également quantifié l'expression de *SNCA* chez les cas sporadiques, les patients SCA2p et les SCA2c ce qui a révélé une sous-expression qui tend à être significative chez les sujets SCA2p et les sporadiques corroborant la diminution observée sur les puces de ces mêmes sujets.

L'ensemble de ces données suggèrent que des perturbations du métabolisme des ARNs sont communes à l'ensemble des sujets étudiés mais que celles-ci impliqueraient des gènes spécifiques selon le phénotype des malades. De façon intéressante, nous avons observé la variation d'expression de *PABPC4* et de *SNCA* chez les parkinsoniens sporadiques. Sachant que *PABPC4* est un paralogue de *PABCP1* appartenant à la famille des protéines PABPs (polyA Binding Protein) et qu'il code une protéine impliquée dans la dégradation des ARNs, nous pouvons émettre l'hypothèse que des perturbations des mécanismes de dégradation pourraient être à l'origine de la diminution de l'expression de *SNCA* et qu'ils pourraient être liés à la physiopathologie de la MP. Des études supplémentaires pour investiguer le rôle de protéines PABPs sont nécessaires pour évaluer la pertinence de cette hypothèse.

[Article 2: RNA binding disturbances as a continuum from Spinocerebellar ataxia type 2 to Parkinson's disease](#)

## RNA binding disturbances as a continuum from Spinocerebellar ataxia type 2 to Parkinson's disease

Aurore Nkiliza<sup>a,b+</sup>, Eugénie Mutez<sup>a,b,c+</sup>, Clémence Simonin<sup>a,b,c+</sup>, Frederic Leprêtre<sup>a,d</sup>, Aurélie Duflot<sup>a,b</sup>, Martin Figeac<sup>a,d</sup>, Alexandre Genet<sup>a,e</sup>, Bernard Sablonnière<sup>a,e</sup>, David Devos<sup>a,c</sup>, Luc Defebvre<sup>a,c</sup>, Alain Destée<sup>a,b,c</sup>, Marie-Christine Chartier-Harlin<sup>a,b,\*</sup>

<sup>a</sup> Université de Lille, F-59000 Lille, France

<sup>b</sup> UMR-S 1172 INSERM Jean-Pierre Aubert Research Center, Team “Early stages of Parkinson’s disease”, F-59045 Lille, France

<sup>c</sup> Movement Disorder Department, Neurologic Clinic, University Hospital, F-59000 Lille, France

<sup>d</sup> Functional Genomic Platform, F-59045 Lille, France

<sup>e</sup> Biological Center, CHRU, F-59045 Lille, France

+ Equally contributed to this work

\* Corresponding author at: IRCL, JPArc, Place de Verdun, F-59045 Lille Cedex, France. Tel.: +33 0320 169 222; fax: +33 0320 169 229. E-mail address: [marie-christine.chartier-harlin@inserm.fr](mailto:marie-christine.chartier-harlin@inserm.fr) (M.-C. Chartier-Harlin)

**Keywords:** Parkinson’s disease, Spinocerebellar ataxia type 2, ATXN2, CAG triplets expansion, RNA metabolism



## **Abstract**

CAG expansions in *ATXN2* known as causative for Spinocerebellar ataxia type 2 (SCA2c patients), also lead to inherited forms of Parkinson's disease (PD, SCA2p patients). Repeat interruptions by CAA triplets are common in physiological and pathological alleles in SCA2p, while expanded alleles in SCA2c contain a pure repeat tract. Even though a number of investigations have been launched against size and structure of triplet repeats, little is known on the molecular mechanisms underlying PD phenotype of SCA2p. Here we studied transcriptome profiles of peripheral blood mononuclear cells from 7 SCA2p patients, 6 SCA2c patients, 7 sporadic PD patients and 13 sex and age-matched controls to highlight differences between SCA2c and PD patients. We revealed similarities between the SCA2p and sporadic PD groups in genes and pathways already known to be altered in PD but also put forward RNA metabolism disturbances. Similarities between our data and RNA metabolism genes whose expression is dependent of ataxin-2 binding protein PABCP family suggest that the control of translation initiation, nonsense mediated RNA Decay (NMD) and RNA splicing might explain some aspect of PD phenotype in SCA2p patients and distinguish them from SCA2c patients. These findings open up the way to further investigations of PABCP family involvement and the regulation of translation in the broad sense in PD. A better understanding of these cellular processes as pathological events of PD might provide more insights in mechanism contributing to the disease and help to find specific biomarkers.

## 1. Introduction

Parkinson's disease (PD) is a multi-factorial disease with a complex etiology including environmental and genetic factors such as *SNCA*, *LRRK2*, *PRKN*, *DJ-1*, or *PINK1*. With the information obtained from GWAS nowadays, such genetic influence has been established to reach at least 27% and more than 33% in some population [1]. Interestingly, the study of inherited forms pinpointed the involvement of several cellular processes in PD onset that have also been found in sporadic PD [2]–[4]. Thus, inherited forms of PD provided valuable insight into the molecular mechanisms underlying its pathophysiology.

Besides those well recognized PD genes, others genes have been linked to Parkinsonism such *ATXN2* initially known as causative gene of Spinocerebellar ataxia type 2 for SCA2c cerebellar patients. This gene is characterized by repeats of CAG triplets in exon 1 whose increased number is pathological. Physiological allele contains 12 to 31 CAG repeats. Expansions of these trinucleotides up to 40 repeats were broadly linked to a cerebellar phenotype while intermediate repeats is now considered as a risk factor for amyotrophic lateral sclerosis (ALS) and/or Parkinsonism [5]. Furthermore, CAG triplet expansions in *ATXN2* from 29 to 32 have also been described as responsible for late onset L-DOPA responsive PD for SCA2p parkinsonian patients, which can segregate as an autosomal dominant inherited forms of PD (SCA2p) although the exact mechanisms by which *ATXN2* triplet expansions lead to PD still need to be better understood.

*ATXN2* encodes the cytoplasmic protein ataxin-2 with physiological function related to several steps of RNA metabolism. It has been shown that ataxin-2 localized in endoplasmic reticulum, stress granules and P-bodies, all involved in the regulation of gene expression, RNA degradation and RNA translation. Ataxin-2 comprised different functional domain. The PAM2 domain mediated its interaction with the cytoplasmic poly(A) binding protein PABCP1, which is involved in translation initiation, nonsense-mediated mRNA decay and RNA splicing [6]–[8]. The Lsm/Lsm-AD domains traditionally carried by RNA splicing factors [6], [9] sustained its implication in RNA splicing together with the ataxin-2 interaction with RBM9, RBPMS, TDP-43 and FUS, four RNA splicing factors [7]. Finally, ataxin-2 has been recently described as a RNA binding protein interacting with 3'UTR of mRNA to regulate their stability [8]. Thus, several lines of evidence emphasized the involvement of ataxin-2 in several processes related to "RNA processing" assuming that disturbances of "RNA metabolism" might occur in patients with *ATXN2* mutations.

It has been proposed that SCA2p may differ from SCA2c by size and structure of CAG repeats. It could thus be queried whether these differences may or not affect ataxin-2 functions in the same manner and whether the related disturbances would be responsible for the different phenotypes.

Peripheral blood mononuclear cells (PBMCs) constitute an easily accessible window onto the brain transcriptome [10]. Knowing that our previous transcriptome analyses of PBMCs revealed perturbations in pathways known to be involved in PD-related neurodegeneration, in the present study, we performed a transcriptome profiling analysis from PBMCs of SCA2c, SCA2p and also sporadic PD *versus* controls. We highlighted changes in gene expression involved in RNA metabolism both in SCA2c, SCA2p and sporadic PD patients. Some of them were specific to patients with a PD phenotype bringing regulation of RNA splicing and translation mechanisms to the forefront in PD.

## **2. Materiel and methods**

### *2.1. Subjects*

Participants were recruited at the Movement Disorder Unit of the Lille University Hospital. Patients enrolled in this study gave their written informed consent to participate to the study, which was approved by the local ethics committee. All patients were thoroughly examined by a neurologist specialized in movement disorders. Blood samples were obtained from 13 patients heterozygous for *ATXN2* mutation (7 females and 6 males), 7 sporadic PD patients (4 females, 3 males) and 13 unrelated, gender and age matched-controls (7 females and 6 males). Seven patients heterozygous for *ATXN2* mutation with a mean age of  $66.1 \pm 12.7$  years were classified as SCA2p patients because of their clinically typical PD. Six other with a mean age of  $47.3 \pm 13.8$  years were classified as cerebellar SCA2c patients since they presented clinical manifestations of Spinocerebellar ataxia type 2. One patient also had a moderate parkinsonian syndrome, a classical observation in Spinocerebellar ataxia type 2. Seven sporadic PD patients without *ATXN2* mutations and with a mean age of  $65.2 \pm 9.9$  years old were diagnosed with probable PD according international criteria [11]. The 7 SCA2p and sporadic PD patients were compared to the same aged and gender matched controls with a mean age of  $65.9 \pm 12.4$  years old, while SCA2c were compared to their respective gender and aged matched controls with a mean age of  $52.8 \pm 8.4$  years old and were free of any personal or familial neurological disorders.

### *2.2. Sequencing*

*ATXN2* gene was amplified by PCR using the following amplification cycles: 1' at 94°C, 14' at 64°C, 1' at 72°C with primers GGGCCCTCACCATGTTCG and CCAATGTCCGCAAGCCCG. The amplicons were sequenced on a 3130XL Genetic Analyser (Applied Biosystems). Both alleles were sequenced together, and then, in order to determine on which allele were the heterozygous CAA triplets, should this happen, the normal allele was sequenced again.

### 2.3. RNA extraction

Blood samples were collected from each subject in BD Vacutainer<sup>®</sup> CPT Cell Preparation Tube containing sodium heparin (Le Pont-de-Claix, France) and RNA were extracted as described in [4]. RNA integrity number (RIN) was checked by using the RNA 6000 NanoChip kit on an Agilent 2100 Bioanalyzer (Agilent Technologies, Massy, France). All RNA samples with RIN >8 were further analyzed by microarrays.

### 2.4. Microarray procedure

The procedure used Agilent<sup>®</sup> one-color whole human genome 44K microarrays according to manufacturer's instructions. Steps and analysis were performed as described in [4]. Raw microarray data were deposited in the Gene Expression Omnibus (GEO) database, [www.ncbi.nlm.nih.gov/geo](http://www.ncbi.nlm.nih.gov/geo) (accession number GSE ). Three lists of transcripts significantly deregulated (Welch's t-test, cut-off  $p < 0.05$  and a fold change of 1.3) were generated corresponding to 2326 probes in SCA2c, 1553 in SCA2p and 1877 in PD either to 1720, 1061 and 1320 unique genes respectively. In order to obtain significantly associated canonical pathways and define ontology and related biological processes, we subjected differentially expressed genes to Ingenuity Pathways Analysis IPA<sup>®</sup> ([www.ingenuity.com](http://www.ingenuity.com), with cut-off  $p < 0.05$ ) and GSEA software (<http://www.broadinstitute.org/gsea/index.jsp>, with cut-off  $q < 0.05$ ) [12]. Using Webgestalt (<http://bioinfo.vanderbilt.edu/webgestalt/>) pathways from these 3 lists were compared with KEGG pathways of publicly available data of mRNA targeted by ataxin-2 and ataxin-2  $\Delta$ PAM2, which lacked the ability to bind to PABPC1 (accession number of the data bank of Japan under DRA001158). These comparisons enabled to test the potential targets of ataxin-2 linked or not to its PAM2 domain in SCA2p, SCA2c and sporadic PD groups [8]. Detailed RNA metabolism genes were investigated by using Toppgene (<https://toppgene.cchmc.org/>). Finally, we compared our data with dataset corresponding to a whole-genome expression array (GeneChip Human Genome U 133 Plus 2.0 arrays, Affymetrix) analysis from *SNpc* of parkinsonian PD at different Braak stages (GEO accession number GSE49036; Dijkstra et al., 2015) as well as our previous PBMCs transcriptome profiles from PD patients versus controls (Whole Human Genome Microarray 4x44K G4112F, GEO accession number GSE22491 and GSE49126). We re-analyzed the data with the GEO2R web tool (<http://www.ncbi.nlm.nih.gov/geo/geo2r/>) using the Linear Models For Microarray Analysis (LIMMA) package in R software, in order to identify genes that were expressed differentially in each of the two datasets. The threshold for statistical significance was set to  $p < 0.05$ .

### *2.5. Quantitative real-time Reverse Transcription Polymerization Chain Reaction (RT-qPCR)*

Validations of selected differentially expressed genes were performed by RT-qPCR with StepOnePlus Real-Time PCR System (Applied Biosystems) using the samples used in microarrays experiments and primers listed in table S1. As far as possible, primers were designed to target Agilent probes to validate gene expression dysregulations. *MAN2B1* was used as endogenous control in all experiments. Analysis of relative gene expression data was performed using the  $2^{-\Delta\Delta CT}$  method [13]. Statistical analysis of qPCR data were performed by using non parametric Mann-Whitney test. Data are presented as mean  $\pm$  SEM. Significant differences were indicated with asterisks considering  $p < 0.05$  as significant threshold.

## **3. Results**

### *3.1. Difference of CAG repeats configuration in SCA2c and SCA2p*

We sequenced *ATXN2* in order to determine the number and the configuration of nucleotide repeats in each group. All sporadic PD patients presented a normal number of repeats at the limit of pathological range (31 triplets). Unlike sporadic PD, SCA2c and SCA2p showed a greater number of repeats than 31 triplets. There were a slightly longer expansions in cerebellar (37-45 triplets) than in parkinsonian patients (35-39).

No CAA interruptions were observed in subjects presenting the longest CAG expansions (45 triples) and a single CAA interruption among other SCA2c patients. In contrast, in SCA2p, the number of CAA interruptions seemed to be related to the lengths of CAG repeats present on the disease allele. No interruption was present in the subjects having the shorter mutated alleles (32 and 35 triplets) while 4 to 5 CAA interruptions existed for those having the longer mutated alleles (37 and 39 triplets) (data not shown).

### *3.2. Specific canonical pathways to each patients group*

To better understand molecular events linked to the 3 groups, canonical pathways were analyzed. While in the 3 groups, some pathways such as Immune system and PI3K/Akt/mTOR/eIF4/p70S6K signaling appeared as central mechanisms of this neurodegenerative disorder and were previously found in familial and sporadic PD patients [4], [14]–[17], others were more specific to each group (Table 1A). Canonical pathways specific to each group were listed in Table S2A, S2B and S2C. SCA2c patients were characterized by perturbations of different pathways related to Huntington's disease signaling, inositol phosphate metabolism and cancer. These

observations were interesting since *ATXN2* mutations leading to SCA2c have been described as a reducer of proapoptotic activity in neuroblastoma cells [18] and as an inducer of  $\text{Ca}^{2+}$  release by interacting with inositol (1,4,5)-trisphosphate receptors [19]–[21].

In turn, SCA2p patients were characterized by dysregulation of pathways involved in the establishment of contacts between neighboring cells (epithelial adherens junction signaling, remodeling of epithelial adherens junctions, germ cell-Sertoli cell Junction signaling) and cytoskeleton (ILK signaling) but also by several pathways regulating cell growth, proliferation and differentiation such as TGF- $\beta$  signaling, ERK5 signaling, HIF1 $\alpha$  signaling and human embryonic stem cell pluripotency. These might be related to the known functions of ataxin-2 described with an abundant expression in embryonic and adult tissues and to ataxin-2-binding protein 1 (A2BP1) shown to promote cell differentiation [22].

Remarkably, sporadic PD patients displayed specific dysregulation in accordance with their phenotype such as variations of Parkinson's disease signaling, cell death, oxidative stress, mitochondrial dysfunctions, immunity and inflammation, Unfolded Protein Response, Endoplasmic Reticulum stress pathway and EIF2 signaling; all these mechanisms being already associated with PD physiopathology [4].

### *3.3. Resemblances between SCA2p and sporadic PD groups*

To test whether differences in size of CAG repeat might explain molecular changes responsible for developing either PD phenotype or SCA2c, we searched for pathways only disturbed in SCA2p and sporadic PD patients and not in SCA2c. We found perturbations of dTMP de novo biosynthesis, Natural killer cell signaling, Ephrin receptor signaling, Actin cytoskeleton signaling, Wnt/ $\beta$ -catenin signaling and Leukocyte extravasation signaling (Table 1B). Of note, these last 4 pathways were also disturbed in PBMCs from *LRRK2* patients. Interestingly, Amyotrophic lateral sclerosis (ALS) signaling was perturbed in both SCA2p and sporadic PD suggesting the involvement of potential common molecular mechanisms. Knowing the important role of RNA metabolism in ALS, one could think it might also be involved in patients with a PD phenotype, ie in SCA2p and sporadic PD patients. Of note, our previous transcriptome study of PD patients heterozygous for a *LRRK2* mutation emphasized expression variations of genes involved in RNA processing and RNA translation. Altogether, these 3 different transcriptome data tended to demonstrate the implication of these processes in PD [2], [4]. Strikingly, we also observed that the SCA2p group shared a higher number of deregulated genes with sporadic PD patients compared to those in the SCA2c group (Figure 1). Among genes dysregulated in both SCA2p and sporadic PD patients, some were already reported by us and others as dysregulated in PD [23], [24] (Table 2A) including *SNCA* and *ALAS2* while others

including *TARDBP* and *DDX6* were found as responsible for ALS or linked to RNA metabolism (Table 2B).

All together, these results suggested a link between those different forms of PD and RNA metabolism.

#### 3.4. Differences in RNA metabolism between SCA2c and SCA2p

Since RNA metabolism is known to be disturbed in SCA2c, we wonder whether such dysfunction would be observed in the pathways analyses data. As expected, “cleavage and polyadenylation of pre-mRNA” and “p70S6K signaling” known to regulate translation were specifically perturbed in SCA2c group (Table S2A). Together with the aforementioned data, all of these results suggested that the “RNA metabolism” was disturbed in SCA2c, SCA2p and sporadic patients.

Knowing that the phenotypes of SCA2p and SCA2c patients resulted from mutations in the same gene, we hypothesized that these mutations might differently affect the functions of ataxin-2 in the “RNA metabolism”, leading thus to phenotypic differences. Only few data regarding the role on RNA metabolism related to specific domain of *ATXN2* are available in the literature. Of interest, Yokoshi *et al.* characterized the mRNA targets interacting with ataxin-2 depleted or not of this PAM2 domain which enabled the binding to PABPC1. Thus, to investigate whether the functions of ataxin-2 might differ depending on its interaction with PABPC1, we selected the genes specific of the PAM2 domain or not. KEGG pathways analyses obtained with the specific targets revealed certain pathways disturbances linked to Parkinson's disease as well as several aspects of RNA metabolism (mRNA surveillance pathway, spliceosome, RNA degradation, Ribosome biogenesis in eukaryotes) and ALS signaling. “Parkinson's disease” and “mRNA surveillance pathway” were both obtained in the lists of ataxin-2 targets related PABCP1 and those of ataxin-2 depleted of its PAM2 domain (data not shown), while “spliceosome”, “RNA degradation”, “Ribosome biogenesis in eukaryotes” and “Amyotrophic lateral sclerosis (ALS)” were deregulated only when PABPC1 could interact with ataxin-2 (Table S3), suggesting that this protein may modulate the functions of ataxin-2. Interestingly enough, analyses of canonical pathways from deregulated genes in SCA2p, SCA2c and sporadic groups have shown that the “Amyotrophic lateral sclerosis (ALS)” was deregulated in the groups of Parkinson phenotypes (SCA2p and sporadic) but not in SCA2c group. The interaction of PABPC1 to ataxin-2 being associated with the perturbation of “Amyotrophic lateral sclerosis (ALS)”, it might be possible that changes involving PABPC1 would be one of the cause leading to differences in phenotypes between parkinsonian and SCA2c.

#### 3.5. Specificity of RNA metabolism disturbances

Given the role of PABPC1 in RNA metabolism, we further investigated perturbations of such cellular processes in SCA2c and SCA2p as well as sporadic PD patients. Analyses by GSEA software identified significant enrichments of deregulated genes related to this metabolism in SCA2c and SCA2p patients (Table 3A and 3B). Such perturbations also occurred in sporadic PD patients in agreement with the aforementioned RNA metabolism involvement in PD (Table 3C). In all 3 groups of patients, “RNA metabolism” defects entailed up and downregulated genes with a lower proportion of downregulated genes in SCA2p (Figure 2). When looking for the most affected “RNA metabolism” process in these groups, the “RNA binding in polyA tail” was among the highest disturbed molecular functions in SCA2c ( $p=2.91E-53$ ), in SCA2p ( $p=6.87E-37$ ) and sporadic PD ( $p=4.49E-45$ ) (Figure 3). “Translation factor activity, RNA binding” related to Translation seemed to be more perturbed in SCA2p ( $p=5.46E-07$ ) and sporadic PD ( $p=3.17E-11$ ) groups. SCA2p and SCA2c presented common impairment of “mRNA 3’UTR binding” ( $p=7.68E-11$  in SCA2c and  $p=1.28E-07$  in SCA2p). This information was interesting considering that ataxin-2 has been shown as a regulator of mRNA stability by interacting with 3’UTR [8]. Finally, we observed that sporadic PD might differentiate from the other groups by perturbation of RNA splicing (Figure 3A, 3B and 3C).

Regarding the common impairment of RNA binding, we hypothesized that it could involve specific genes underlying different phenotypes. Four genes deregulated either in PD patients or in SCA2c on our microarray data were selected to confirm results by RT-qPCR (Figure 4). Because of the major role of  $\alpha$ -synuclein in PD as well as hnRNPs and MBNL in splicing, we also quantified the *SNCA*, *hnRNPUL1*, *hnRNPA3*, *TARDBP* and *MBNL1* genes which were reported to be deregulated on our microarrays. RT-qPCR data revealed a significant downregulation of *hnRNPA3*, *hnRNPUL1* and *MBNL1* in SCA2c patients as shown in microarrays. Significant downregulation of *MBNL1* was observed in sporadic PD patients on microarrays and qPCR data, but. Opposite direction of expression was observed for *PABPC4* qPCR’s data. However we also noted a significant downregulation of this gene in our previous transcriptome analyses of PBMCs from 18 sporadic PD versus 20 controls as well in those from 10 LRRK2 patients versus 8 controls [2] (Agilent probe A\_23\_P51718;  $p<0.05$  FC 0.80 and A\_23\_P51718  $p=7.02E-08$  FC=0.54, respectively). In addition in the latter, PABPC1 probes was also downregulated (Agilent probes A\_23\_P82693  $p=9.05E-04$ , FC=0,78) [4]. We thus also analyzed other microarrays data to evaluate whether PABPC1 in SNpc would be deregulated in sporadic PD patients compared to controls. Interestingly, we observed a significant 2.27 fold decrease of *PABPC1* in SNpc of sporadic PD (Affymetrix probe 215157xat;  $p=0.01$ ). In accordance with microarrays data, *SNCA* expression was strongly decreased in SCA2p and sporadic PD patients while no such difference was observed in SCA2c patients (Table 4).



## 2. Discussion

Growing evidences emphasize the interest in studying familial forms of PD to understand molecular mechanisms underlying this disease's phenotype. Here, we focused on *ATXN2*, a gene in which CAG triplet expansions have been associated with classical PD phenotype in familial PD. By using PBMCs of patients heterozygous for *ATXN2* mutations, we confirmed that differences in length of CAG triplets expansions lead to gene expression differences which can be specific of each phenotype, cerebellar or PD phenotype. The transcriptome comparisons obtained with *ex vivo* cells from sporadic PD patients and their respective controls revealed similarities in terms of genes and pathways altered both in SCA2p and sporadic PD patients that might contribute to the occurrence of the same phenotype. In addition, our data highlighted the involvement of specific RNA metabolism perturbations in the different parkinsonian phenotypes and also classical spinocerebellar ataxia.

### 2.1. Similarity between SCA2p and sporadic PD patients

Several studies emphasized existence of the same clinical phenotype between patients with heterozygous expansion of CAG triplet *in ATXN2* and sporadic PD patients without providing any explanation for this similarity. Our results offered some potential clarification by showing higher overlap between SCA2p and sporadic PD patients in terms of deregulated genes. Among these, we confirmed deregulations of genes linked to pathways already associated to PD such as heme biosynthesis and iron metabolism. Interestingly increased level of iron in *SNpc* has been associated to PD by generating many reactive oxygen species responsible for neuronal death. Reducing the level of iron in *the SNpc* by the use of chelating seemed to be beneficial on PD symptoms in patients and is currently being tested in clinical trials sustaining the importance of iron in PD [25]. Here, we displayed deregulation of *ALAS2*, *HBD* and *HBM*, genes involved in such pathways, exclusively in patients showing a PD phenotype, both in SCA2p and sporadic PD groups. These genes were previously described as being deregulated *in SNpc* [26] but also in our previous PBMCs transcriptome studies of sporadic PD and heterozygous G2019S *LRRK2* patients. Giving that a growing body of evidences related the involvement of the immune response and lymphocyte *SNpc* infiltration in PD [3], [4], [17], [26], [27] it was also interesting to note that transcripts involved in leukocyte extravasation signaling were commonly altered in sporadic and *ATXN2* PD patients. The SCA2p also resembled to sporadic PD patients by the deregulation of genes involved in Wnt/ $\beta$ -catenin and ALS signaling; these might be sustained by the overlap of repeats number in ALS and Parkinsonism [5]. The latter comprises *BCL2*, *BCL2L1*, *CACNA1E* and *CACNA1C* genes all involved in regulation of

intracellular calcium level and cell death through disturbances of mitochondria. Furthermore, a recent study defined that the risk for ALS linked to repeats of CAG in the *ATXN2* increases began with 29 repeats and reached a maximum at 32 and 33 repeat trinucleotide repeat [28], a disease well known for its link with impairment of RNA metabolism. This disease was characterized by the presence of cytoplasmic proteinaceous aggregates rich in TDP-43, a nuclear RNA binding protein (RBP) having several implications in pre-mRNA splicing, mRNA stability and microRNA metabolism. Repeats expansions in *C9orf72* and *ATXN2* genes have been described as modifiers of TDP-43 toxicity acting on its aggregation, mislocation and sequestration. Moreover, mutations in *TARDBP* gene as well as *FUS*, *TAF15*, *hnRNPA1/B2*, *hnRNPA1* and *hnRNPA3* were also reported to be responsible for rare causes of ALS [29]–[31] and it should be noted that the expression of some of these genes were deregulated in our transcriptome data. All of these proteins were known to be part of hnRNP family linked to processing heterogeneous nuclear RNAs (hnRNAs) into mature messenger RNAs (mRNAs) through biological functions including mRNA export, localization, translation and stability. Thus, identification of ALS signaling deregulation in sporadic PD patients and SCA2p patients might pointed out the involvement of RNA metabolism in the development of a parkinsonian phenotype.

## 2.2. RNA metabolism perturbations both in heterozygous *ATXN2* patients and sporadic PD patients

Our data tend to confirm that the distinction between SCA2p and SCA2c might not to lie in size but in the structure of repeats since CAG repeat numbers has been shown to overlap between these two groups [32]. Sobczak *et al.* pointed out that CAG repeats in *ATXN2* could induce specific three dimensional structures such as hairpins [33]. These hairpin loops would be defined by the number of repeats and/or CAA interruptions number and location. In several triplet repeat disorders, these RNA structures have been shown to allow sequestration of proteins involved in gene expression via microRNA, RNA splicing and RNA translation strengthening the hypothesis that such cellular processes could contribute to the development of these disorders. RNA metabolism seems to be important both for SCA2p and SCA2c patients. This is confirmed by our RT-qPCR data measuring expression of genes, all members of RNA binding protein family except for the  $\alpha$ -synuclein. These proteins were characterized by the presence of RNA-binding domains (RBDs) in single or multiple copies through which they interacted directly with RNA molecules. RBPs were implicated in a broad range of functions including posttranscriptional regulation of mRNA stability, splicing, editing and translation, mRNA export and localization, mRNA polyadenylation and miRNA biogenesis. All these observations were in favor of a role of RNA in SCA2c and parkinsonian phenotypes. However, even if gene alterations related to RNA metabolism were observed both in SCA2c, SCA2p and also in sporadic PD groups, our comparisons of transcriptome profile and qPCR data suggest a specificity of

these alterations regarding the patient groups or their phenotype. Thus, some specific RNA metabolism signature might be sufficient to distinguish SCA2c and PD phenotypes and to identify pathophysiological specificities of these two disorders.

### 2.3. *ALS signaling discrepancy in PD patients highlight a potential involvement of stress granule in PD*

Our data have highlighted a number of signaling pathways similarities between the SCA2p and sporadic groups including deregulation of genes involved in ALS signaling pathway. We hypothesized that those gene variations were linked to triplet expansion in the *ATXN2* in the SCA2p group. Surprisingly, while CAG expansions are also responsible for cerebellar phenotype in SCA2c, the ALS signaling pathway was observed in this group. To further investigate these data, we focused on genes whose expression has been shown to be modulated by ataxin-2 by Yokoshi *et al.* [8]. We observed that the deregulation of genes associated with the “ALS signaling pathway” only occurred when the PAM2 domain of ataxin-2 was present (Table S3A and S3B). This domain has been shown to mediate the interaction of ataxin-2 with PABPC1, a protein involved in nonsense-mediated mRNA decay, RNA splicing and translation initiation [34]–[36]. Of interest PABPC1 has been shown to amplify the TDP-43 toxicity in yeast and *Drosophila* models of ALS but also the toxicity induced by CAG triplet expansion in *ATXN2* with 32 trinucleotide expansions. This protein also contributed to translation repression by inhibiting the phosphorylation of eIF2 $\alpha$  protein and promoting stress granules formation where untranslated RNAs are stored. We did not observed any transcriptome changes concerning *PABPC1* expression in PBMCs. We noted a significant decrease of *PABPC4* expression, a paralog of *PABPC1*, in the microarray data from SCA2p and sporadic PD patients. However, we did not replicate these results by PCR, probably because we could not quantitate the same transcripts in the microarrays and the qPCR. Indeed, when looking for the *PABPC4* expression in previous PBMCs transcriptome data from PD versus Controls, a decrease of expression in Patients was systematically observed. In addition, the majority of the selected genes for qPCR experiments as well as for *CAPZA2* gene not related to RNA metabolism, have shown gene variations in the same direction as for the microarray data.. Interestingly, a specific search for expression variation of *PABPC 1* revealed its deregulation in the *SNpc* of PD patients compared to controls. Unlike PABPC1 which was expressed in most cells, PABPC4 was found to be expressed in activated T-cells [37]. PABPC4 shared a high sequence homology with PABPC1 with the same functional domains including its interaction with RNAs to regulate their translation. Thus, in PBMCs the overall translational regulation properties of PABPC4 might distinguish SCA2c from patients with PD phenotype. Ataxin-2 has also been shown to interact with the DDX6 protein, a helicase which like PABPC1 was located in stress granules. DDX6 has been reported to be associated with several component of miRNA

biogenesis machinery such as Ago1, Ago2, and the CCR4-NOT in drosophila. This protein has also been shown to localize and associated with mRNA degradation proteins DCP2, DCP1 and EDC3 in processing bodies containing many enzymes for storage in regulating mRNA, translation, and degradation. Strikingly, we observed an overexpression of DDX6 in both the SCA2p and sporadic PD groups and a downregulation in the SCA2c group (Table 2). Altogether these data supported the involvement of proteins of stress granules and translation in molecular changes given rise to parkinsonian or cerebellar phenotypes.

#### 2.4. Conclusion

To our knowledge, this is the first transcriptome study investigating the participation of *ATXN2* mutations in PD phenotype. By analyzing the different profiles from PD patients with or without CAG expansions in *ATXN2*, we highlighted i) a shared deregulation of pathways that have already been described in PD cases, ii) disturbances of RNA metabolism in SCA2c, SCA2p and PD PBMCs revealing a putative continuum from Parkinson's disease to Spinocerebellar ataxia type 2. While ataxin-2 has diverse role in RNA metabolism, our data suggest that this metabolism might make difference between PD phenotype and SCA2c phenotype. This seems to require involvement of PABP family by interaction with ataxin-2 and focus on translation initiation, NMD and RNA splicing significance in PD. Further investigations are required to evaluate the full scope of RNA metabolism in PD and the role of PABP proteins in PD onset.

#### **Conflict of interest**

None of the authors has a conflict of interest to declare in relation to the present research.

#### **Acknowledgements**

We would like to thank all of the subjects who donated their time and biological samples to be a part of this study as well as Neurological Clinic Department for their important contribution to this work and Yokoshi *et al.* for their help in providing gene lists targeted by ataxin-2[8]. This work was supported by University of Lille, INSERM-French Research Ministry, the Lille University Hospital and the Institut de Recherche sur le Cancer de Lille. The authors also thank Dr. David Fraser for advice on the manuscript's English.

#### **References**

- [1] M. F. Keller, M. Saad, J. Bras, *et al.*, "Using genome-wide complex trait analysis to quantify missing heritability' in Parkinson's disease". *Hum. Mol. Genet.* 2012;21:4996–5009

- [2] E. Mutez, L. Larvor, F. Leprêtre, *et al.*, “Transcriptional profile of Parkinson blood mononuclear cells with LRRK2 mutation”. *Neurobiol. Aging*. 2011;32:1839–1848
- [3] D. T. Dexter and P. Jenner, “Parkinson disease: from pathology to molecular disease mechanisms”. *Free Radic. Biol. Med.* 2013;62:132-44
- [4] E. Mutez, A. Nkiliza, K. Belarbi, *et al.*, “Involvement of the immune system, endocytosis and EIF2 signaling in both genetically determined and sporadic forms of Parkinson’s disease”. *Neurobiol. Dis.* 2014;63:165-170
- [5] C. Gellera, N. Ticozzi, V. Pensato, *et al.*, “ATAXIN2 CAG-repeat length in Italian patients with amyotrophic lateral sclerosis: risk factor or variant phenotype? Implication for genetic testing and counseling”. *Neurobiol. Aging*. 2012;33:1847.e15–21
- [6] U. Nonhoff, M. Ralser, F. Welzel, *et al.*, “Ataxin-2 interacts with the DEAD/H-box RNA helicase DDX6 and interferes with P-bodies and stress granules”. *Molecular biology of the cell*. 2007;18:1385–1396
- [7] J. J. Magaña, L. Velázquez-Pérez, and B. Cisneros, “Spinocerebellar ataxia type 2: Clinical presentation, molecular mechanisms, and therapeutic perspectives”. *Molecular Neurobiology*. 2013;47:90–104
- [8] M. Yokoshi, Q. Li, M. Yamamoto, *et al.*, “Direct Binding of Ataxin-2 to Distinct Elements in 3’ UTRs Promotes mRNA Stability and Protein Expression.” *Mol. Cell*. 2014;55:186–198
- [9] C. Kaehler, J. Isensee, U. Nonhoff, *et al.*, “Ataxin-2-like is a regulator of stress granules and processing bodies” *PLoS One*. 2012;7:e50134
- [10] I. S. Kohane and V. I. Valtchinov, “Quantifying the white blood cell transcriptome as an accessible window to the multiorgan transcriptome”. *Bioinformatics*. 2012;28: 538–45
- [11] G. S. Gelb DJ1, Oliver E, “Diagnostic Criteria for Parkinson Disease”. *Arch Neurol*. 1999;56:33–39
- [12] A. Subramanian, P. Tamayo, V. K. Mootha, *et al.*, “Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles”. *Proc. Natl. Acad. Sci. U. S. A.* 2005;102:15545–50

- [13] K. J. Livak and T. D. Schmittgen, "Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method". *Methods*. 2001;25:402–408
- [14] J. Infante, C. Prieto, M. Sierra, *et al.* "Identification of candidate genes for Parkinson's disease through blood transcriptome analysis in LRRK2-G2019S carriers, idiopathic cases, and controls". *Neurobiol. Aging*. 2015;36:1105–1109
- [15] V. Grozdanov, C. Bliederaeuser, W. P. Ruf, *et al.*, "Inflammatory dysregulation of blood monocytes in Parkinson's disease patients," *Acta Neuropathol*. 2014;128:651–663
- [16] E. Grünblatt, S. Mandel, J. Jacob-Hirsch, *et al.*, "Gene expression profiling of parkinsonian substantia nigra pars compacta; alterations in ubiquitin-proteasome, heat shock protein, iron and oxidative stress regulated proteins, cell adhesion/cellular matrix and vesicle trafficking genes". *J. Neural Transm*. 2004;11:1543–1573
- [17] M. D. Chikina, C. P. Gerald, X. Li, *et al.*, "Low-Variance RNAs Identify Parkinson's Disease Molecular Signature in Blood". *Mov. Disord*. 2015;30:813-21
- [18] R. Wiedemeyer, F. Westermann, I. Wittke, *et al.*, "Ataxin-2 promotes apoptosis of human neuroblastoma cells". *Oncogene*. 2003;22:401–411
- [19] J. Liu, T.-S. Tang, H. Tu, *et al.*, "Deranged Calcium Signaling and Neurodegeneration in Spinocerebellar Ataxia Type 2". *J. Neurosci*. 2009;29:9148–9162
- [20] I. Bezprozvanny, "Role of inositol 1,4,5-trisphosphate receptors in pathogenesis of Huntington's disease and spinocerebellar ataxias" *Neurochem. Res*. 2011;36:1186–1197
- [21] D. V. Adebimpe W. Kasumi, Xia Liang, *et al.*, "Chronic suppression of inositol 1,4,5-triphosphate receptor-mediated calcium signaling in cerebellar Purkinje cells alleviates pathological phenotype in spinocerebellar ataxia 2 mice". *J Neurol Sci*. 2012;32:12786–12796
- [22] O. Y. Tastan, J. Z. Maines, Y. Li, *et al.*, "Drosophila Ataxin 2-binding protein 1 marks an intermediate step in the molecular differentiation of female germline cysts". *Development*. 2010;137:3167–3176
- [23] M. Chesselet and B. Ritz, "Transcriptional regulation of  $\alpha$ -synuclein: insights from blood?". *Futur. Neurol*. 2009;4:145–147

- [24] C. R. Scherzer, A. C. Eklund, L. J. Morse, *et al.*, “Molecular markers of early Parkinson’s disease based on gene expression in blood”. *Proc Natl Acad Sci U S A*. 2007;104:955–960
- [25] G. Grolez, C. Moreau, B. Sablonnière, *et al.*, “Ceruloplasmin activity and iron chelation treatment of patients with Parkinson’s disease”. *BMC Neurol*. 2015;15:74
- [26] S. Mandel, “Gene Expression Profiling of Sporadic Parkinson’s Disease Substantia Nigra Pars Compacta Reveals Impairment of Ubiquitin-Proteasome Subunits, SKP1A, Aldehyde Dehydrogenase, and Chaperone HSC-70”. *Ann. N. Y. Acad. Sci*. 2005;1053:356–375
- [27] A. A. Dijkstra, A. Ingrassia, R. X. de Menezes, *et al.*, “Evidence for Immune Response, Axonal Dysfunction and Reduced Endocytosis in the Substantia Nigra in Early Stage Parkinson’s Disease”. 2015;10:e0128651
- [28] A. G. Neuenschwander, K. K. Thai, K. P. Figueroa, *et al.*, “Amyotrophic Lateral Sclerosis Risk for Spinocerebellar Ataxia Type 2 *ATXN2* CAG Repeat Alleles”. *JAMA Neurol*. 2014;71: 1529
- [29] D. Calini, L. Corrado, R. Del Bo, *et al.*, “Analysis of hnRNPA1, A2/B1, and A3 genes in patients with amyotrophic lateral sclerosis”. *Neurobiol. Aging*. 2013;34:2695.e11–2695.e12
- [30] S. J. Barmada, “Linking RNA Dysfunction and Neurodegeneration in Amyotrophic Lateral Sclerosis”. *Neurotherapeutics*. 2015;12:340–351
- [31] J. Cooper-Knock, J. J. Bury, P. R. Heath, *et al.*, “C9ORF72 GGGGCC Expanded Repeats Produce Splicing Dysregulation which Correlates with Disease Severity in Amyotrophic Lateral Sclerosis”. *PLoS One*. 2015;10:e0127376
- [32] P. Charles, A. Camuzat, N. Benammar, F. Sellal, A. Deste, I. Le Ber, and G. Stevanin, “Are interrupted SCA2 CAG repeat expansions responsible for parkinsonism?”. 2007;69:1970-5.
- [33] K. Sobczak and W. J. Krzyzosiak, “CAG repeats containing CAA interruptions form branched hairpin structures in spinocerebellar ataxia type 2 transcripts”. *J. Biol. Chem*. 2005;280:3898–910
- [34] I. Peixeiro, Â. Inácio, C. Barbosa, *et al.*, “Interaction of PABPC1 with the translation initiation complex is critical to the NMD resistance of AUG-proximal nonsense mutations”. *Nucleic Acids Res*. 2012;40:1160–1173

- [35] J. Zhu, H. Ding, X. Wang, *et al.*, "PABPC1 exerts carcinogenesis in gastric carcinoma by targeting miR-34c". *Int. J. Clin. Exp. Pathol.* 2015;8:3794–802
- [36] R. Joncourt, A. B. Eberle, S. C. Rufener, *et al.*, "Eukaryotic Initiation Factor 4G Suppresses Nonsense-Mediated mRNA Decay by Two Genetically Separable Mechanisms". *PLoS One.* 2014;9:e104391
- [37] I. Peixeiro, A. Inacio, C. Barbosa, *et al.*, "Interaction of PABPC1 with the translation initiation complex is critical to the NMD resistance of AUG-proximal nonsense mutations". *Nucleic Acids Res.* 2012;40:1160–1173

#### A. COMMON PATHWAYS IN THE 3 GROUPS

Canonical Pathways	Sporadic PD patients	SCA2p patients	SCA2c patients
	<i>p</i> -value	<i>p</i> -value	<i>p</i> -value
Primary Immunodeficiency Signaling	1.05E-02	2.69E-03	1.78E-07
Regulation of eIF4 and p70S6K Signaling	1.20E-02	4.07E-02	3.16E-02
B Cell Development	1.32E-02	2.00E-02	8.32E-04
ERK/MAPK Signaling	1.41E-02	1.45E-02	2.40E-03
Docosahexaenoic Acid (DHA) Signaling	2.51E-02	3.39E-02	2.40E-03
mTOR Signaling	2.82E-02	3.09E-02	2.63E-03
PI3K/AKT Signaling	2.88E-02	1.26E-02	1.32E-02

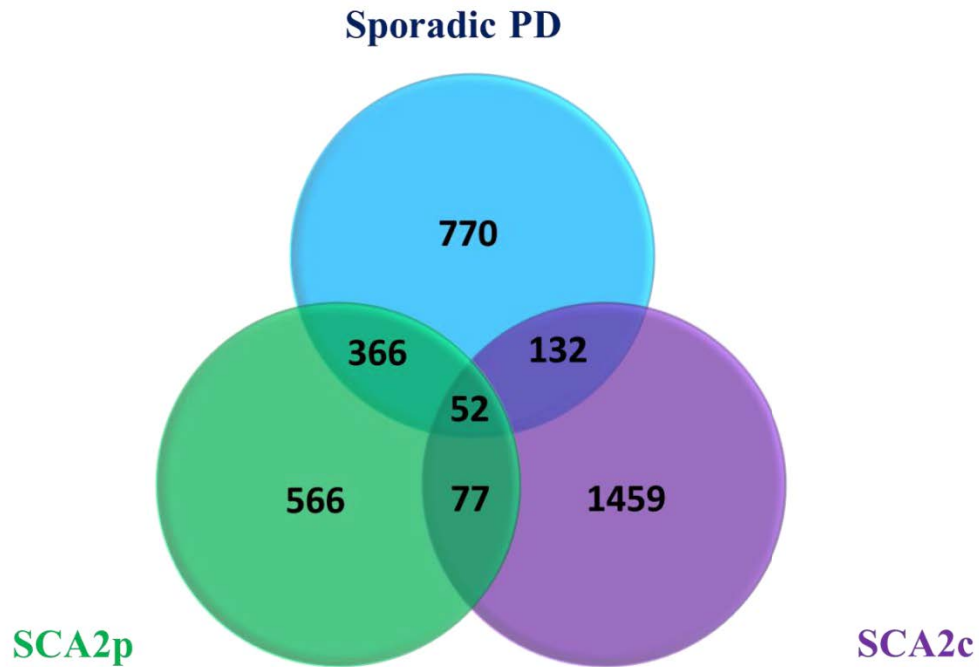
#### B. COMMON PATHWAYS BETWEEN SCA2p AND SPORADIC PD

Tec Kinase Signaling	2.34E-03	1.62E-02
Leukocyte Extravasation Signaling	3.09E-03	1.00E-03
Ephrin Receptor Signaling	3.16E-03	3.39E-02
Wnt/ $\beta$ -catenin Signaling	5.01E-03	1.29E-02
Amyotrophic Lateral Sclerosis Signaling	1.23E-02	6.03E-03
Natural Killer Cell Signaling	1.23E-02	1.38E-02
Actin Cytoskeleton Signaling	1.70E-02	2.63E-02
dTMP De Novo Biosynthesis	3.09E-02	2.00E-02

**Table 1: Significantly deregulated pathways identified by the Ingenuity Pathways Analysis (IPA)**

Fischer's exact test was used to determining the probability of the association between deregulated genes in all of the 3 patient groups (1A) or all PD patients (1B) and canonical pathway ( $p < 0.05$ ).





**Figure 1: Venn diagram representation of differentially expressed genes in each patient group.**

The number of overlapping genes between SCA2p and sporadic PD groups was higher compared to the other comparisons.

Gene Symbol	Entrez Gene Name	SCA2c		SCA2p		Sporadic PD	
		FC	Agilent Probe	FC	Agilent Probe	FC	Agilent Probe
<b>A. PD-RELATED GENES</b>							
<b>ALAS2</b>	Aminolevulinatase, delta-, synthase 2		N.D	-8.50	A_32_P22654	-17.92	A_32_P22654
				-8.41	A_32_P385587	-23.44	A_32_P385587
<b>HBD</b>	Hemoglobin, delta		N.D	-8.46	A_23_P87346	-8.42	A_23_P87346
				-5.84	A_24_P75190	-10.36	A_24_P75190
<b>HBM</b>	Hemoglobin, mu		N.D	-4.34	A_23_P15055	-6.59	A_23_P15055
<b>SELENBP1</b>	Selenium binding protein 1		N.D	-4.90	A_23_P86021	-13.03	A_23_P86021
<b>SNCA</b>	Synuclein, Alpha	1.44	A_23_P29939	-2.01	A_32_P109653	-1.77	A_32_P109653
<b>B. ALS-RELATED GENES</b>							
<b>BCL2</b>	B-Cell CLL/Lymphoma 2	-1.38	A_23_P352266	2.20	A_23_P208132	2.35	A_23_P208132
<b>BCL2L1</b>	BCL2-Like 1		N.D	-1.48	A_23_P210886	-1.43	A_23_P210886
<b>CACNA1E</b>	Calcium Channel, Voltage-Dependent, R Type, Alpha 1E Subunit		N.D	2.66	A_23_P34554	3.06	A_23_P34554
<b>CACNA1C</b>	Calcium Channel, Voltage-Dependent, L Type, Alpha 1C Subunit		N.D	1.35	A_23_P373031	1.33	A_23_P373031
<b>TARDBP</b>	TAR DNA Binding Protein	1.45	A_32_P234935		N.D		N.D
		1.32	A_24_P276703				
<b>DDX6</b>	DEAD (Asp-Glu-Ala-Asp) Box Helicase 6	-1.47	A_32_P80255	1.31	A_24_P215407	1.37	A_24_P215407
<b>hnRNPA1</b>	Heterogeneous Nuclear Ribonucleoprotein A1		N.D		N.D	-1.50	A_32_P189803
<b>hnRNPA3</b>	Heterogeneous Nuclear Ribonucleoprotein A3	-1.36	A_24_P488586		N.D	-1.61	A_24_P269895

**Table 2: Selection of PD and ALS-related genes in each group**

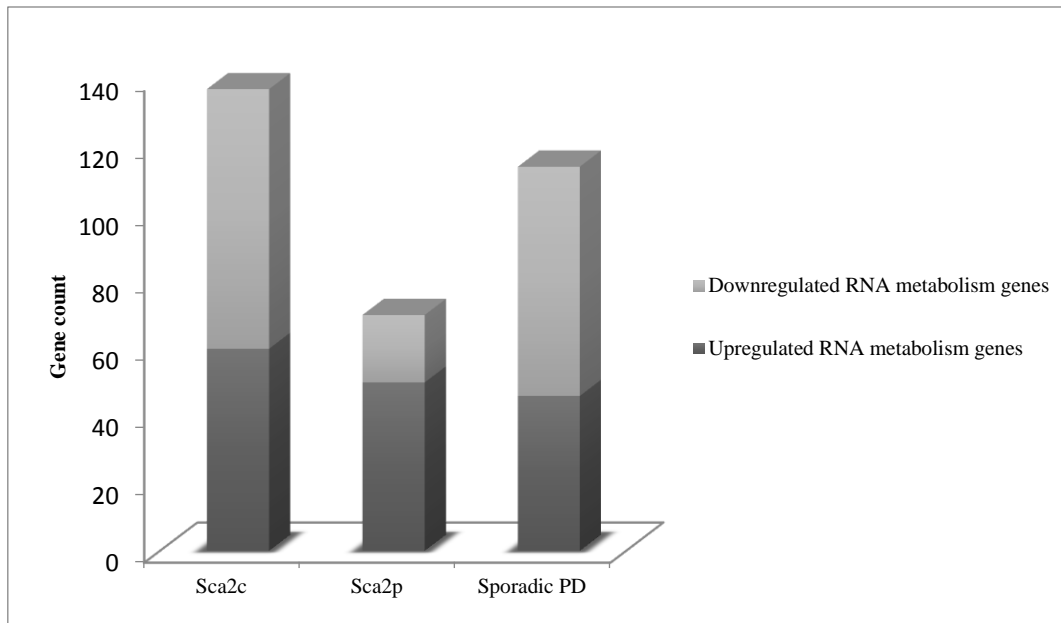
FC: Fold change showing the gene expression level associated to the Agilent probes.

ND: not deregulated

GO terms	Genes (%)	<i>p-value</i>	FDR <i>q-value</i>	Description
<b>A. SCA2c</b>				
GO:0043283	13.7%	4.28E-71	3.53E-68	biopolymer metabolic process
GO:0006139	13.6%	1.55E-51	6.4E-49	nucleobasenucleosidenucleotide and nucleic acid metabolic process
GO:0016070	14.6%	7.59E-41	2.09E-38	RNA metabolic process
GO:0019222	14.1%	3.47E-36	7.15E-34	regulation of metabolic process
GO:0044260	11.9%	4.74E-35	7.82E-33	cellular macromolecule metabolic process
GO:0031323	14.0%	8.93E-35	1.23E-32	regulation of cellular metabolic process
GO:0044267	11.8%	7.8E-34	9.2E-32	cellular protein metabolic process
<b>B. SCA2p</b>				
GO:0043283	6.5%	5.82E-25	4.8E-22	biopolymer metabolic process
GO:0048856	7.6%	1.77E-21	7.32E-19	anatomical structure development
GO:0006139	6.5%	1.73E-18	4.75E-16	nucleobasenucleosidenucleotide and nucleic acid metabolic process
GO:0016070	7.6%	4.44E-18	9.16E-16	RNA metabolic process
GO:0007275	6.9%	9.34E-18	1.54E-15	multicellular organismal development
GO:0007165	5.6%	8.37E-17	1.15E-14	signal transduction
GO:0019538	6.3%	1.15E-16	1.36E-14	protein metabolic process
<b>C. Sporadic PD</b>				
GO:0043283	9.3%	1.89E-42	1.56E-39	biopolymer metabolic process
GO:0006139	10.1%	5.5E-37	2.27E-34	nucleobasenucleosidenucleotide and nucleic acid metabolic process
GO:0016070	10.9%	3.92E-30	1.08E-27	RNA metabolic process
GO:0007165	7.7%	3.25E-26	6.71E-24	signal transduction
GO:0006350	10.5%	7.82E-25	1.29E-22	transcription
GO:0006351	10.9%	1.18E-22	1.46E-20	transcription DNA dependent
GO:0019538	8.1%	1.36E-22	1.46E-20	protein metabolic process

**Table 3: Top biological processes enriched in each group obtained with GSEA**

The top 7 GO terms related to deregulated genes in SCA2c (3A), SCA2p (3B) and sporadic PD patients (3C) are listed in the table.



**Figure 2: Proportion of RNA metabolism genes up and downregulated in each group**

Number of RNA metabolism genes was lower in SCA2p compared to the two other groups. The percentages of downregulated genes represented 56%, 29% and 60% in SCA2c, SCAp and sporadic PD respectively.

Figure 3A. SCA2c



Figure 3B. SCA2p

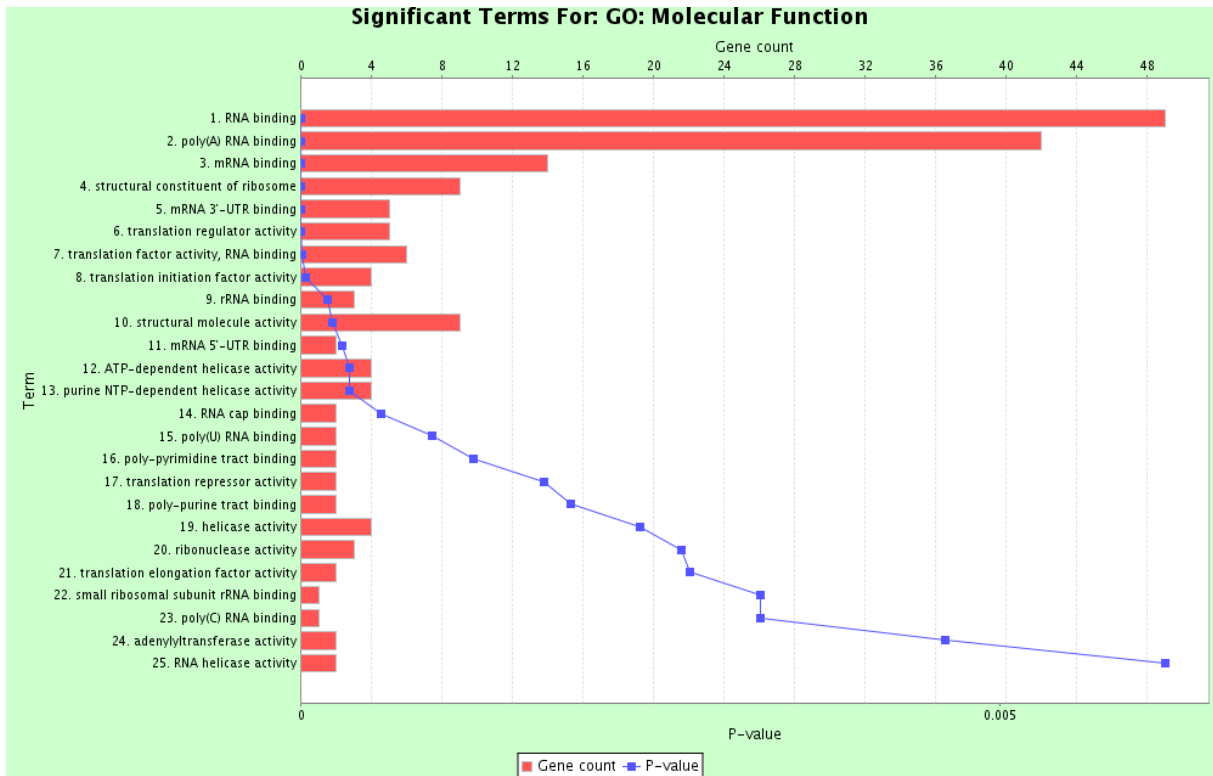
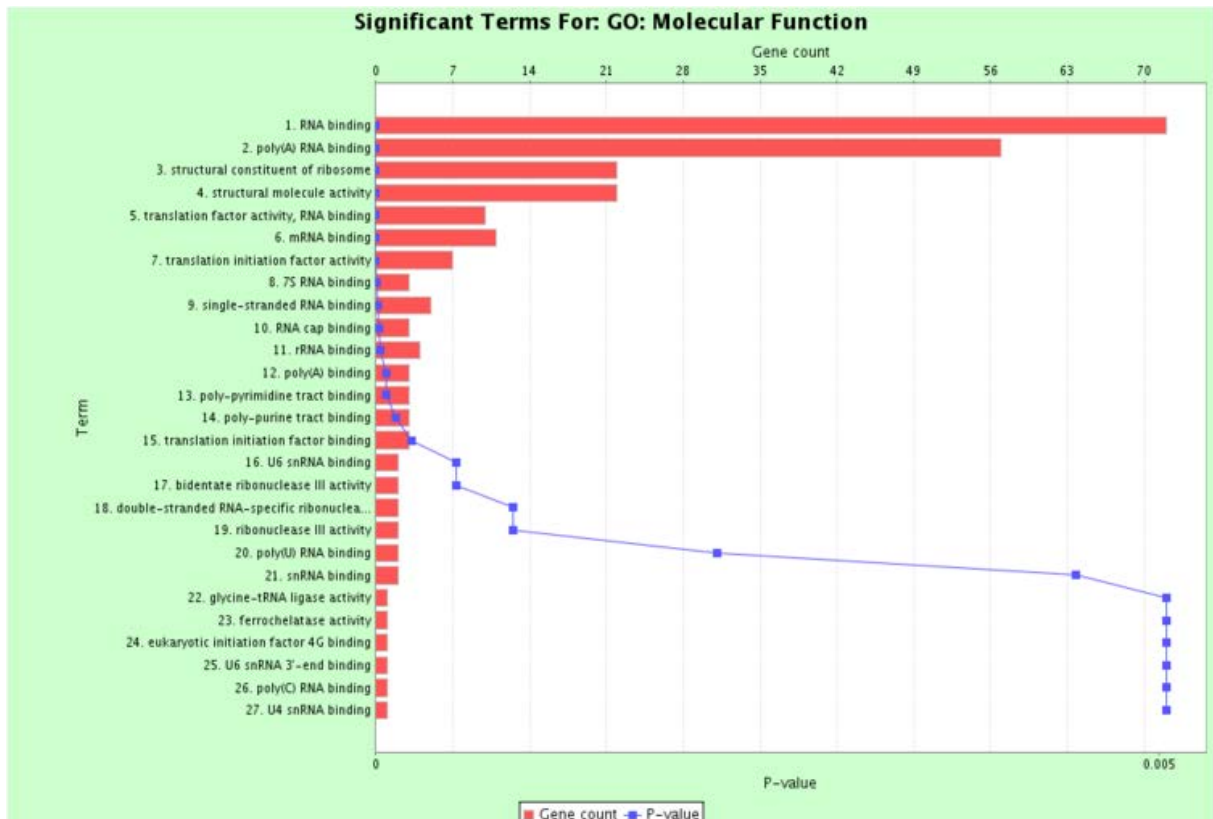
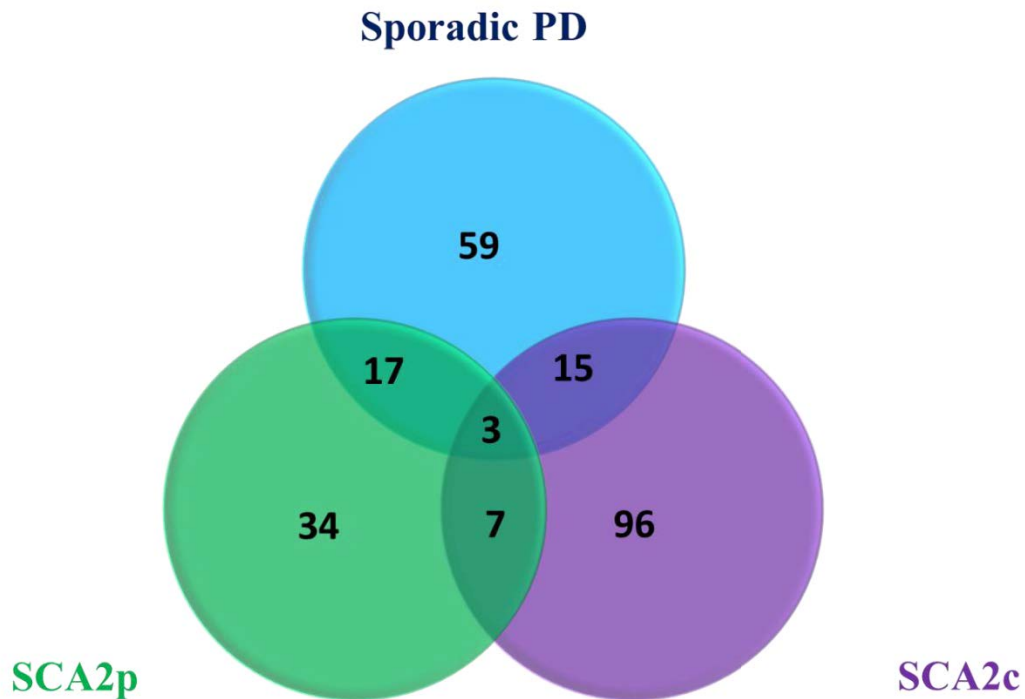


Figure 3C. Sporadic PD



**Figure 3: Molecular functions related to RNA metabolism in each group**



**Figure 4: Venn diagram representation of differentially expressed RNA metabolism genes**

RNA metabolism genes were reported as deregulated in all the 3 groups with some overlap between different patients groups. Dysregulated probes corresponding to *PABPC4*, *RPL22* and *RPL9* were the same in SCA2p and sporadic PD while deregulated probes corresponding to *hnRNPA3* in SCA2c and sporadic PD were different.

Agilent Probes	Gene Name	FC microarrays			FC qPCR		
		SCA2c	SCA2p	Sporadic PD	SCA2c	SCA2p	Sporadic PD
A_23_P307940	CAPZA2		*-1,35	*-1,72	-1,22	-1,20	*-2,58
A_24_P488586	hnRNPA3	*-1.36			*-1.54	-1.07	-1.35
A_24_P108351	hnRNPUL1	*-1.31			*-1.28	1.04	-1.14
A_24_P925361	MBNL1	*-1.47			*-1.69	-1.44	*-2.46
A_23_P51718	PABPC4		*-1.30	*-1.42	-1.38	1.01	*2.10
A_23_P29939/A_32_P109653	SNCA total	*1.46/N.D	*-2.12/*-2.01	*-2.12/*-1.77	1.16	-2.05	-2.19
A_24_P276703	TARDBP	*1.32			-1.35	-1.15	-1.10

**Table 4: Validations of microarray results by RT-qPCR**

Agilent probes were targeted by primers to validate gene expression deregulations. \* $p < 0.05$

Primer Name	Primers Sequences	Melting temperature (°C)	Amplicon length (bp)
hnRNPA3-Forward	ACTGTTTATTGGTGGTCTG	57	83
hnRNPA3-Reverse	AATCTGTGAGTGTGCCC		
hnRNPUL1-Forward	AACCAGTATCAGCAGTATG	55	87
hnRNPUL1-Reverse	GTAGTCGTAGTTCCTCGTAG		
MBNL1-Forward	TTGTATTCAGTCTTGTGGCT	57	134
MBNL1-Reverse	AATGATCTAGCGGTTAATGG		
SNCA-Forward	ACAGCAGTAGCCCAGAAGACAG	60	198
SNCA-Reverse	GGCTTCAGGTTTCGTAGTCTTGATA		
PABPC4-Forward	GGCAGGCAGAGTAAAACGG	60	172
PABPC4-Reverse	TCCTCCAGCATTACCTTAGCAC		
TARDBP-Forward	TCATCACGCATCACAGG	58	157
TARDBP-Reverse	ACCATCAATCCCACACAAG		

**Table S1: Primers used for microarray validations by RT-qPCR**

Detailed RT-qPCR conditions were as followed: cDNA synthesis was made from 250ng of total RNA by using High-Capacity cDNA Reverse Transcription (Life technologies) as instructed by the supplier. qPCR were performed by using Power SYBR® Green Master Mix (Life technologies). All assays were run in triplicate. The qPCR conditions were 50 °C for 2min and 95 °C for 10min followed by 10 cycles of touchdown (95°C for 15 sec, 10°C above annealing temperature with touchdown -1°C each cycle) to conclude by 40 cycles at 95 °C for 15 sec, annealing temperature for 60 sec with primers in the table above.

**Table S2A. Specific to SCA2c**

Canonical Pathways (IPA; p<0.05)	p-value
Prostate Cancer Signaling	4.17E-05
Glioblastoma Multiforme Signaling	1.26E-04
Myc Mediated Apoptosis Signaling	3.80E-04
Chronic Myeloid Leukemia Signaling	6.76E-04
p70S6K Signaling	7.41E-04
Non-Small Cell Lung Cancer Signaling	1.20E-03
Telomerase Signaling	1.41E-03
TREM1 Signaling	1.58E-03
Superpathway of Inositol Phosphate Compounds	1.70E-03



Gap Junction Signaling	1.78E-03
<b>HER-2 Signaling in Breast Cancer</b>	1.82E-03
<b>Glioma Signaling</b>	2.34E-03
GM-CSF Signaling	2.45E-03
<b>Small Cell Lung Cancer Signaling</b>	2.75E-03
FcγRIIB Signaling in B Lymphocytes	3.47E-03
3-phosphoinositide Biosynthesis	3.89E-03
<b>Colorectal Cancer Metastasis Signaling</b>	4.68E-03
Production of Nitric Oxide and Reactive Oxygen Species in Macrophages	5.89E-03
Role of p14/p19ARF in Tumor Suppression	6.76E-03
CTLA4 Signaling in Cytotoxic T Lymphocytes	7.24E-03
<b>Estrogen-Dependent Breast Cancer Signaling</b>	7.24E-03
Cell Cycle: G1/S Checkpoint Regulation	9.12E-03
Sphingosine-1-phosphate Signaling	9.33E-03
<b>Cleavage and Polyadenylation of Pre-mRNA</b>	1.05E-02
FLT3 Signaling in Hematopoietic Progenitor Cells	1.05E-02
HMGB1 Signaling	1.05E-02
Leptin Signaling in Obesity	1.05E-02
Circadian Rhythm Signaling	1.15E-02
Melanocyte Development and Pigmentation Signaling	1.17E-02
Thrombin Signaling	1.20E-02
FGF Signaling	1.29E-02
Melanoma Signaling	1.35E-02
Amyloid Processing	1.48E-02
RAR Activation	1.66E-02
NGF Signaling	1.70E-02
UVA-Induced MAPK Signaling	1.70E-02
<b>Huntington's Disease Signaling</b>	1.86E-02
IL-2 Signaling	1.91E-02
IL-3 Signaling	1.95E-02
P2Y Purigenic Receptor Signaling Pathway	2.04E-02
Notch Signaling	2.45E-02
Phosphatidylethanolamine Biosynthesis II	2.69E-02
Angiopoietin Signaling	2.88E-02
IL-15 Signaling	2.88E-02
Insulin Receptor Signaling	2.88E-02
TR/RXR Activation	2.95E-02
Erythropoietin Signaling	3.16E-02
Role of PKR in Interferon Induction and Antiviral Response	3.16E-02
Branched-chain α-keto acid Dehydrogenase Complex	3.24E-02
<b>Myo-inositol Biosynthesis</b>	3.24E-02
<b>Superpathway of D-myo-inositol (1,4,5)-trisphosphate Metabolism</b>	3.31E-02
Gαq Signaling	3.39E-02
<b>D-myo-inositol (1,4,5,6)-Tetrakisphosphate Biosynthesis</b>	3.47E-02
<b>D-myo-inositol (3,4,5,6)-tetrakisphosphate Biosynthesis</b>	3.47E-02

IGF-1 Signaling	3.47E-02
Macropinocytosis Signaling	3.47E-02
Role of MAPK Signaling in the Pathogenesis of Influenza	3.47E-02
<b>Breast Cancer Regulation by Stathmin1</b>	3.80E-02
iCOS-iCOSL Signaling in T Helper Cells	3.80E-02
ErbB4 Signaling	3.89E-02
CREB Signaling in Neurons	3.98E-02
Synaptic Long Term Potentiation	3.98E-02
CXCR4 Signaling	4.47E-02
<b>Endometrial Cancer Signaling</b>	4.47E-02
<b>Ovarian Cancer Signaling</b>	4.47E-02
Antiproliferative Role of TOB in T Cell Signaling	4.57E-02
Regulation of the Epithelial-Mesenchymal Transition Pathway	4.57E-02
Cell Cycle Regulation by BTG Family Proteins	4.90E-02
<b>D-myo-inositol-5-phosphate Metabolism</b>	4.90E-02
JAK/Stat Signaling	4.90E-02

**Table S2B. Specific to SCA2p**

Canonical Pathways (IPA; p<0.05)	p-value
<b>Epithelial Adherens Junction Signaling</b>	1.62E-04
<b>ERK5 Signaling</b>	2.51E-03
Remodeling of Epithelial Adherens Junctions	4.27E-03
PRPP Biosynthesis I	6.31E-03
<b>Germ Cell-Sertoli Cell Junction Signaling</b>	1.82E-02
<b>Granulocyte Adhesion and Diapedesis</b>	1.91E-02
<b>Sertoli Cell-Sertoli Cell Junction Signaling</b>	2.00E-02
<b>TGF-<math>\beta</math> Signaling</b>	2.04E-02
<b>ILK Signaling</b>	2.82E-02
Superpathway of Serine and Glycine Biosynthesis I	3.89E-02
Methylthiopropionate Biosynthesis	4.68E-02
<b>HIF1<math>\alpha</math> Signaling</b>	4.90E-02
<b>Human Embryonic Stem Cell Pluripotency</b>	4.90E-02

**Table S2C. Specific to Sporadic PD**

Canonical Pathways (IPA; p<0.05)	p-value
<b>EIF2 Signaling</b>	3.89E-06
Mechanisms of Viral Exit from Host Cells	5.13E-04
Agrin Interactions at Neuromuscular Junction	6.31E-04
CCR5 Signaling in Macrophages	6.31E-04
<b>Death Receptor Signaling</b>	9.33E-04
Virus Entry via Endocytic Pathways	2.04E-03
PEDF Signaling	2.75E-03

L-cysteine Degradation III	3.47E-03
Arginine Degradation VI (Arginase 2 Pathway)	3.55E-03
Aryl Hydrocarbon Receptor Signaling	3.55E-03
Proline Biosynthesis II (from Arginine)	3.55E-03
<b>Calcium-induced T Lymphocyte Apoptosis</b>	3.98E-03
<b>Apoptosis Signaling</b>	5.75E-03
Signaling by Rho Family GTPases	5.75E-03
<b>Mitochondrial Dysfunction</b>	6.03E-03
Communication between Innate and Adaptive Immune Cells	6.92E-03
Rac Signaling	7.94E-03
Axonal Guidance Signaling	8.91E-03
<b>Ephrin B Signaling</b>	1.00E-02
T Cell Receptor Signaling	1.12E-02
<b>Parkinson's Signaling</b>	1.23E-02
Heme Biosynthesis II	1.29E-02
Pathogenesis of Multiple Sclerosis	1.29E-02
<b>Unfolded protein response</b>	1.29E-02
Role of CHK Proteins in Cell Cycle Checkpoint Control	1.45E-02
RAN Signaling	1.51E-02
Dendritic Cell Maturation	1.82E-02
<b>Differential Regulation of Cytokine Production in Macrophages and T Helper Cells by IL-17A and IL-17F</b>	1.86E-02
Proline Biosynthesis I	1.91E-02
Rapoport-Luebering Glycolytic Shunt	1.91E-02
Antigen Presentation Pathway	2.00E-02
TNFR1 Signaling	2.34E-02
Basal Cell Carcinoma Signaling	2.45E-02
NF-κB Activation by Viruses	2.69E-02
Prolactin Signaling	2.69E-02
<b>Xenobiotic Metabolism Signaling</b>	2.88E-02
Nitric Oxide Signaling in the Cardiovascular System	3.09E-02
PCP pathway	3.09E-02
Tetrapyrrole Biosynthesis II	3.09E-02
α-Adrenergic Signaling	3.09E-02
<b>Endoplasmic Reticulum Stress Pathway</b>	3.16E-02
Cholesterol Biosynthesis I	3.72E-02
Cholesterol Biosynthesis II (via 24,25-dihydrolanosterol)	3.72E-02
Cholesterol Biosynthesis III (via Desmosterol)	3.72E-02
Paxillin Signaling	3.72E-02
<b>eNOS Signaling</b>	3.80E-02
Cyclins and Cell Cycle Regulation	3.89E-02
<b>Regulation of IL-2 Expression in Activated and Anergic T Lymphocytes</b>	4.17E-02
CCR3 Signaling in Eosinophils	4.27E-02
<b>Differential Regulation of Cytokine Production in Intestinal Epithelial Cells by IL-17A and IL-17F</b>	4.27E-02
Glycerol Degradation I	4.47E-02

Fcy Receptor-mediated Phagocytosis in Macrophages and Monocytes	4.68E-02
Growth Hormone Signaling	4.90E-02
RhoGDI Signaling	4.90E-02
Tumoricidal Function of Hepatic Natural Killer Cells	4.90E-02

**Table S3A. Specific KEEG pathways related to mRNA only targeted by ataxin-2**

<b>Specific KEEG pathways of mRNA specifically targeted by ataxin-2 with PAM2 domain</b>	<b>Adjusted p-value</b>
Spliceosome	4.52E-18
Protein processing in endoplasmic reticulum	4.52E-18
Ubiquitin mediated proteolysis	8.84E-15
Wnt signaling pathway	5.77E-14
Pyrimidine metabolism	1.82E-10
DNA replication	6.75E-10
Adherens junction	1.47E-09
RNA degradation	5.05E-09
Regulation of actin cytoskeleton	3.26E-08
N-Glycan biosynthesis	9.86E-08
Renal cell carcinoma	1.27E-07
Nucleotide excision repair	1.28E-07
Vasopressin-regulated water reabsorption	1.28E-07
Focal adhesion	5.97E-07
Basal transcription factors	6.65E-07
ErbB signaling pathway	8.31E-07
Ribosome biogenesis in eukaryotes	9.34E-07
RNA polymerase	2.33E-06
Bacterial invasion of epithelial cells	2.79E-06
Lysosome	3.23E-06
Apoptosis	3.26E-06
Chronic myeloid leukemia	4.72E-06
Endometrial cancer	6.51E-06
Base excision repair	8.56E-06
Shigellosis	9.10E-06
Colorectal cancer	1.11E-05
TGF-beta signaling pathway	3.09E-05
VEGF signaling pathway	3.29E-05
Small cell lung cancer	3.50E-05
Phagosome	4.56E-05
Non-small cell lung cancer	4.59E-05
Glycerophospholipid metabolism	5.87E-05
Axon guidance	8.39E-05

Mismatch repair	1.00E-04
Vibrio cholerae infection	2.00E-04
Long-term potentiation	2.00E-04
Notch signaling pathway	2.00E-04
Glioma	3.00E-04
Inositol phosphate metabolism	3.00E-04
Cell adhesion molecules (CAMs)	3.00E-04
Natural killer cell mediated cytotoxicity	3.00E-04
Fc gamma R-mediated phagocytosis	3.00E-04
Thyroid cancer	6.00E-04
Neurotrophin signaling pathway	6.00E-04
Cysteine and methionine metabolism	6.00E-04
Acute myeloid leukemia	1.20E-03
Glycosylphosphatidylinositol(GPI)-anchor biosynthesis	1.50E-03
Tight junction	2.20E-03
Lysine degradation	2.40E-03
Homologous recombination	2.60E-03
Biosynthesis of unsaturated fatty acids	2.60E-03
Leukocyte transendothelial migration	3.40E-03
B cell receptor signaling pathway	3.60E-03
Mucin type O-Glycan biosynthesis	3.80E-03
Pathogenic Escherichia coli infection	3.80E-03
T cell receptor signaling pathway	4.20E-03
Protein export	4.40E-03
Epithelial cell signaling in Helicobacter pylori infection	4.80E-03
Fc epsilon RI signaling pathway	5.20E-03
Alanine, aspartate and glutamate metabolism	5.20E-03
Hepatitis C	5.40E-03
Melanogenesis	5.50E-03
Bladder cancer	6.40E-03
Chagas disease (American trypanosomiasis)	6.90E-03
mTOR signaling pathway	6.90E-03
<b>Amyotrophic lateral sclerosis (ALS)</b>	<b>7.70E-03</b>
Osteoclast differentiation	7.90E-03
Prion diseases	8.20E-03
Hedgehog signaling pathway	1.07E-02
Viral myocarditis	1.56E-02
Sphingolipid metabolism	1.65E-02
Amoebiasis	1.89E-02
Glycosphingolipid biosynthesis - ganglio series	1.91E-02
Vitamin digestion and absorption	2.28E-02
Glycosaminoglycan biosynthesis - heparan sulfate	3.05E-02
p53 signaling pathway	3.30E-02
Collecting duct acid secretion	3.51E-02
Riboflavin metabolism	4.18E-02
Glycerolipid metabolism	4.61E-02

Glutathione metabolism	4.61E-02
------------------------	----------

**Table S3B. Specific KEEG pathways related to mRNA only targeted by ataxin-2 deleted of PAM2 domain**

<b>Specific KEEG pathways of mRNA specifically targeted by ataxin-2 without PAM2 domain</b>	<b>Adjusted <i>p</i>-value</b>
Galactose metabolism	5.00E-04
Fructose and mannose metabolism	1.00E-03
Porphyrin and chlorophyll metabolism	2.50E-03
Butirosin and neomycin biosynthesis	3.20E-03
Starch and sucrose metabolism	4.10E-03
Fatty acid elongation in mitochondria	6.40E-03
Propanoate metabolism	6.60E-03
Adipocytokine signaling pathway	6.60E-03
Fatty acid metabolism	1.25E-02
Type II diabetes mellitus	1.52E-02
Selenocompound metabolism	1.86E-02
Glyoxylate and dicarboxylate metabolism	1.98E-02
Glycosaminoglycan biosynthesis - chondroitin sulfate	2.68E-02
Aminoacyl-tRNA biosynthesis	2.68E-02
Jak-STAT signaling pathway	4.74E-02

### 3.3. Chapitre 3 : Mise en évidence de perturbations de l'épissage dans la maladie de Parkinson

L'ensemble des données que nous avons acquises jusqu'à présent montre que le métabolisme des ARNs semble être un mécanisme commun aux formes génétiques et aux formes sporadiques de la MP. En effet, notre première analyse avait mis en évidence une perturbation de la voie de signalisation EIF2 chez les sujets porteurs de la mutation G2019S de *LRRK2* et les parkinsoniens sporadiques. Notre seconde étude avait révélé des variations d'expression de gènes impliqués dans différents aspects du métabolisme des ARNs avec des modifications de l'expression de gènes liés à l'épissage chez les parkinsoniens porteurs de la mutation de l'*ATXN2* et chez des cas sporadiques.

A l'aide du logiciel GSEA, nous avons analysé les processus biologiques dans lesquels sont impliqués les gènes dérégulés chez les cas sporadiques, les parkinsoniens porteurs de la mutation G2019S de *LRRK2* ou de l'*ATXN2*. Cette analyse a confirmé l'existence de perturbations du métabolisme des ARNs chez l'ensemble des sujets (Tableau 5).

Processus biologiques	Sporadiques	<i>LRRK2</i>	<i>ATXN2</i>
	FDR <i>q</i> -value	FDR <i>q</i> -value	FDR <i>q</i> -value
Processus des biopolymères métaboliques	1,44E-34	1,15E-41	4,80E-22
Processus métabolique des acides nucléiques et nucleotide nucleoside nucléobase	1,99E-29	9,35E-25	4,75E-16
<b>Processus du métabolisme des ARNs</b>	<b>1,20E-25</b>	<b>1,11E-16</b>	<b>9,16E-16</b>
Processus de biosynthèse des ARNs	3,94E-17	2,56E-12	5,97E-12
Transcription	4,14E-17	6,81E-13	4,84E-13
Processus du métabolisme des protéines	5,79E-17	5,31E-45	1,36E-14
Processus du métabolisme des protéines cellulaires	4,01E-16	6,68E-44	1,20E-11
Processus métaboliques des macromolécules cellulaires	7,03E-16	2,05E-43	2,00E-11
Régulation de l'expression génique	9,15E-16	2,53E-08	9,90E-12
Transduction du signal	2,75E-10	1,13E-20	1,11E-14
Développement des organismes multicellulaires	3,35E-09	6,25E-08	1,54E-15
Régulation positive des processus cellulaires	4,11E-09	2,20E-14	2,56E-12
Régulation positive des processus biologiques	2,17E-08	3,69E-14	5,42E-13
Régulation des processus développementaux	4,38E-06	2,89E-09	2,67E-11

**Tableau 5: Processus biologiques communs aux parkinsoniens sporadiques, aux porteurs de la mutation G2019S de *LRRK2* ou d'expansions de triplets CAG dans *ATXN2* (Analyse GSEA ; FDR<0,05)**

La liste des gènes significativement dérégulés chez les 20 cas sporadiques ( $p < 0,05$  ; Chapitre1), chez les 9 sujets porteurs de la mutation de *LRRK2* ( $p < 0,05$  ; Chapitre1), et chez les 7 sujets présentant une augmentation du nombre de triplets CAG dans *ATXN2* ( $p < 0,05$  ; Chapitre2), ont été soumis à l'analyse par GSEA. Les processus significativement liés aux gènes dérégulés et communs aux 3 groupes de patients sont répertoriés dans le tableau ci-dessus (FDR<0,05). Les perturbations du métabolisme des ARNs sont significativement retrouvées chez l'ensemble des malades (en gras et en rouge)

Afin d'approfondir et de préciser le rôle du métabolisme des ARNs dans la MP, nous avons choisi d'explorer l'épissage dans les cellules mononucléées périphériques sanguines par l'utilisation de la méthode de séquençage des ARNs. Cette étude ayant été initiée dans les débuts de l'utilisation du séquençage des ARNs (RNAseq), nous avons dans un premier temps fait une première analyse pour évaluer la faisabilité et la validité de la technique. À la suite de cette étude, une seconde analyse a été réalisée (étude pilote) pour tester l'hypothèse de l'existence de défauts d'épissage quantitatifs et qualitatifs dans la MP. Cependant, les méthodes de séquençage étant en développement constant, les outils et les génomes de référence que nous avons utilisés pour ces deux analyses ont été différents. Nous avons donc fait une 3<sup>ème</sup> analyse qui, bien qu'elle utilise les données de séquençage obtenues lors des 2 analyses précédentes, nous a permis de limiter les biais



en effectuant une analyse simultanée de l'ensemble des échantillons et en évaluant les défauts d'épissage avec le même outil d'analyse.

### 3.3.1.Méthodes

#### 3.3.1.1. Sujets

Les prélèvements sanguins utilisés pour cette étude proviennent de malades et de témoins identifiés dans le cadre des protocoles PARKFANORD (PARKinson dans les FAMilles du NORD de la France) et CONVERGENCE. Ces deux études ont été approuvées par le Comité de Protection des Personnes du CHRU de Lille et tous les sujets inclus dans l'analyse ont signé un consentement écrit de participation.

Les parkinsoniens choisis pour cette étude ont été sélectionnés selon plusieurs critères : le diagnostic a été posé par un neurologue à l'aide des critères diagnostiques de l'UKPDSBRC (United Kingdom Parkinson's Disease Society Brain Research Center) et des critères du NINDS (National Institute of Neurological Disorders and Stroke)<sup>280</sup> (ANNEXE 4) ; un génotypage a été effectué pour les sujets provenant de familles avec une histoire connue de MP afin d'identifier les mutations génétiques dont ils étaient porteurs ; l'absence de signes cliniques d'autres maladies neurodégénératives a été vérifiée ainsi que la prise de traitements que l'on a choisi de limiter aux traitements anti-parkinsoniens pour minimiser l'effet des traitements sur les variations moléculaires mesurées par l'analyse transcriptomique qui pourraient mettre en évidence des variations sans rapport avec la maladie.

Les témoins recrutés pour cette étude ont été choisis de manière à être appariés en sexe et en âge avec les malades. De façon similaire, les témoins ont été choisis parmi les sujets sains prenant le moins de traitements possible pour éviter toutes interactions pouvant induire des variations qui ne relèveraient pas de la maladie ou de l'absence de maladie.

L'ensemble des sujets dont les échantillons sanguins ont été analysés, sont récapitulés dans les tableaux ci-dessous (tableau 6).

Sujets	Âge (ans)	Sexe	Diagnostic	Forme de la maladie	Evolution (Hoehn et Yahr)	Séquenceur utilisé	N° d'analyse
P227/01	52	Femme	Parkinson	<i>LRRK2</i>		SOLiD 4	1 et 3
P299/01	63	Femme	Parkinson	<i>SNCA</i>		SOLiD 4	1 et 3
P59/04	59	Femme	Parkinson	<i>SNCA</i>	<i>De novo</i>	SOLiD 4	1 et 3
3310121	57	Femme	Témoin			SOLiD 4	1 et 3
3310167	53	Femme	Témoin			SOLiD 4	1 et 3
3310199	64	Femme	Témoin			SOLiD 4	1 et 3
3110025	61	Homme	Parkinson	Sporadique	<i>De novo</i>	Ion Proton	2 et 3
3110147	61	Homme	Parkinson	Sporadique	<i>De novo</i>	Ion Proton	2 et 3
3110150	60	Homme	Parkinson	Sporadique	HY=1	Ion Proton	2 et 3
3110138	59	Femme	Parkinson	Sporadique	HY=1	Ion Proton	2 et 3
3110149	71	Femme	Parkinson	Sporadique	HY=1	Ion Proton	2 et 3
3310250	66	Femme	Témoin			Ion Proton	2 et 3
3310268	57	Femme	Témoin			Ion Proton	2 et 3
3310260	61	Homme	Témoin			Ion Proton	2 et 3
3310245	76	Homme	Témoin			Ion Proton	2 et 3
3310242	61	Homme	Témoin			Ion Proton	2 et 3
3310239	41	Femme	Témoin			Ion Proton	2 et 3
3110028	74	Homme	Parkinson	Sporadique	HY=1,5	SOLiD 4	3
P12/16	38	Femme	Parkinson	<i>LRRK2</i>	<i>De novo</i>	SOLiD 4	3
3110070	65	Femme	Parkinson	Sporadique	HY=2,5	Ion Proton	3
3110159	75	Homme	Parkinson	Sporadique	HY=3	Ion Proton	3
3110162	69	Femme	Parkinson	Sporadique	HY=2,5	Ion Proton	3
3110176	72	Homme	Parkinson	Sporadique	HY=3	Ion Proton	3
3110210	74	Femme	Parkinson	Sporadique	HY=2,5	Ion Proton	3
P30/01	87	Femme	Parkinson	<i>EIF4G1</i>	HY=5	Ion Proton	3
P305/01	71	Femme	Parkinson	<i>SNCA</i>	HY=2	Ion Proton	3
P310/01	38	Femme	Parkinson	<i>LRRK2</i>	HY=2	Ion Proton	3
P353/01	69	Femme	Parkinson	<i>SNCA</i>	HY=5	Ion Proton	3
3310202	69	Homme	Témoin			SOLiD 4	3
3310083	73	Homme	Témoin			Ion Proton	3
3310238	59	Femme	Témoin			Ion Proton	3
3310243	66	Femme	Témoin			Ion Proton	3
3310266	78	Femme	Témoin			Ion Proton	3
3110272	74	Femme	Témoin			Ion Proton	3

**Tableau 6: Sujets recrutés pour l'étude de l'épissage par séquençage des ARNs (n°1 : étude de faisabilité, n°2 : étude pilote, n°3 : analyse combinée)**

Ce tableau récapitule l'ensemble des sujets dont les prélèvements sanguins ont été utilisés pour l'analyse de faisabilité du RNAseq (analyse n°1), l'étude pilote (analyse n°2) et l'analyse combinant les données de l'analyse n°1 et l'analyse n°2 (analyse n°3). Ces sujets incluent des porteurs de la mutation G2019S de *LRRK2* (*LRRK2*), des sujets porteurs de la duplication du locus *SNCA* (*SNCA*), un sujet porteur de la mutation p.R1205H d'*EIF4G1* (*EIF4G1*), des sujets sporadiques et les témoins appariés. Ces sujets incluent des malades à des stades précoces de la maladie (*de novo*, Hoehn et Yahr (HY)  $\leq 1,5$ ) et des sujets à des stades plus tardifs (autres sujets). Les scores HY sont indiqués lorsqu'ils ont pu nous être communiqués par les praticiens. Le type de séquenceur utilisé est également indiqué pour chaque sujet.

Un second set de patients a été prélevé pour valider les données de RNAseq. Ce groupe comprend 19 parkinsoniens (âge moyen :  $61 \pm 2$  ans) et dix-huit témoins (âge moyen :  $56,9 \pm 2,6$ )

### 3.3.1.2. Extraction d'ARNs à partir de cellules mononucléées périphériques sanguines

#### 3.3.1.2.1. Protocole d'extraction des ARNs destinés au séquençage par le séquenceur SOLiD™ 4 « Sequencing by Oligonucleotide Ligation and Detection »

Les prélèvements sanguins ont été effectués dans des tubes CPT® (BD Vacutainer) contenant un gradient de Ficoll qui permet la séparation des cellules mononucléées périphériques sanguines du plasma et autres éléments figurés du sang. Cette séparation se fait par centrifugation à 1750g pendant 20 minutes et à température ambiante. Une fois centrifugés, les lymphocytes et monocytes forment un anneau qui est prélevé par pipetage. Pour éliminer les contaminants éventuels, plusieurs cycles de lavages alternant rinçage au Hank's Balanced Salt Solution (solution saline, Thermo-Fisher) et une centrifugation de 10 minutes ont été effectués. Les cellules sont ensuite comptées puis aliquotées à raison de  $0,5 \cdot 10^7$  cellules dans 700µL de tampon de lyse cellulaire (Buffer RLT, RNeasy Mini Kit Qiagen) auquel 1% de  $\beta$ -mercaptoéthanol a été préalablement ajouté. Ceci permet de lyser les cellules tout en inactivant les RNases (ribonucléases), dénaturées par le  $\beta$ -mercaptoéthanol, facilitant ainsi la récupération des ARNs de bonne qualité (faible taux de dégradation). Une lyse mécanique à l'aide d'un vortex est effectuée afin de maximiser la rupture des membranes cellulaires. Soixante dix pourcent d'éthanol sont ajoutés volume à volume de lysat cellulaire pour précipiter les acides nucléiques. Le mélange est ensuite passé sur colonne de silice qui retient les acides nucléiques et purifie le mélange après 30 secondes de centrifugation à 12000g. Après une étape de lavage, un premier traitement à la DNase (désoxyribonucléase) est effectué pendant 45 minutes à température ambiante pour limiter les contaminations par l'ADN. Une fois l'ADN digéré, plusieurs lavages sont effectués pour éliminer la DNase, les sels apportés par les différents tampons utilisés au cours de l'extraction et les traces d'éthanol. Les ARNs sont ensuite récupérés par élution de la colonne par ajout d'un volume de 30µL d'eau RNase-free (dépourvue de RNase) et une centrifugation de 2 minutes à 12000g. L'éluât est ensuite récupéré pour une seconde digestion à la DNase pour éliminer le maximum d'ADN possible. La suspension d'ARNs a ensuite été séchée au SpeedVac pour concentrer les ARNs en les remettant en suspension dans 10µL. Cette étape de concentration est requise pour la préparation de la librairie qui requiert un kit adapté aux très petits volumes. Une fois l'extraction des ARNs terminée, ceux-ci sont dosés par dosage spectrophotométrique au Nano-Drop® ND-1000 qui mesure l'absorbance à 260nm. Les rapports d'absorbance 260 nm/280 nm et

260 nm/230 nm sont également mesurés pour évaluer la contamination en protéines et en solvant respectivement. Enfin, l'intégrité des ARNs est évaluée grâce au Bioanalyzer RNA 6000 nano d'Agilent© qui détermine le RNA Integrity Number (RIN) pouvant aller de 0 à 10 ; 0 lorsque ARN est complètement dégradé et 10 si l'ARN est intact. Les ARNs utilisés pour notre étude avaient tous un RIN supérieur à 8. Avant d'utiliser les ARNs pour la préparation de la librairie, une PCR de recherche de contamination par l'ADNg est effectuée à l'aide d'amorces oligonucléotides s'appariant avec un intron du gène *LRRK2*. Une quantification de l'expression de *SNCA* est également effectuée pour vérifier que les ARNs soient suffisamment intègres pour permettre leur conversion en ADNg. Les protocoles de PCR et de qPCR sont détaillés paragraphe 3.3.1.8.

#### 3.3.1.2.2. Protocole d'extraction des ARNs destinés au séquençage par le séquenceur Ion Proton™

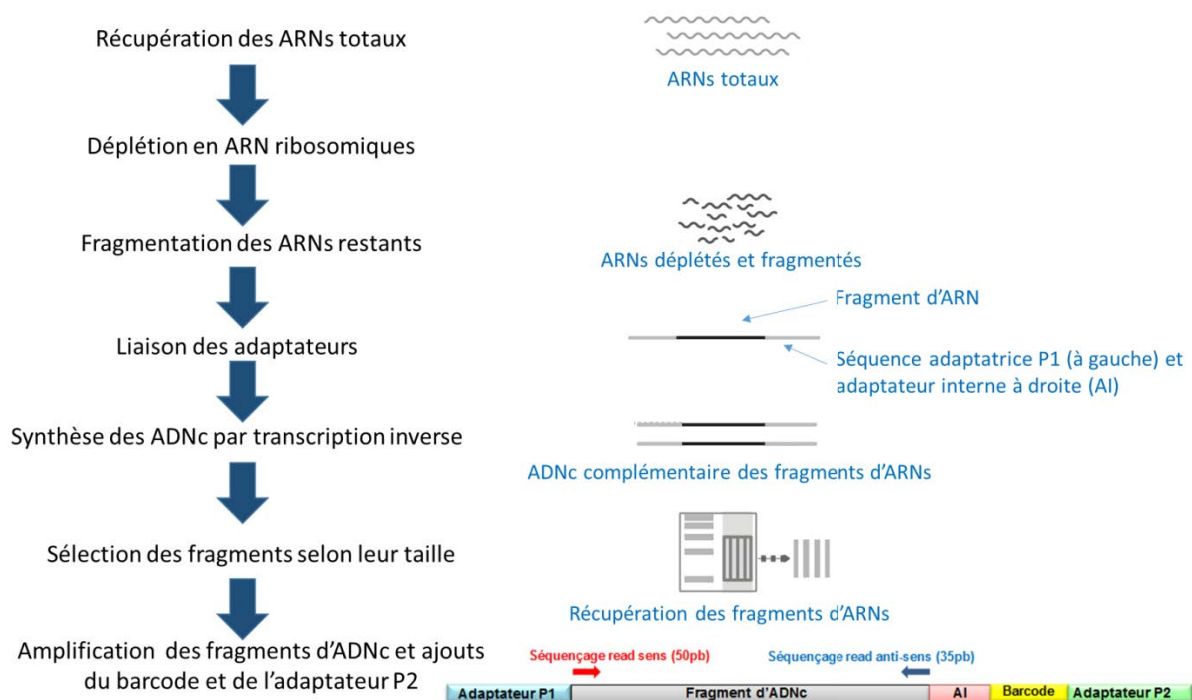
Le protocole suivi pour la récupération des cellules mononucléées, le dosage des ARNs et la mesure de leur intégrité ont strictement été identiques à celui suivi dans le cadre de la procédure SOLiD™ 4. Seul a différé l'extraction des ARNs qui dans ce cas précis a été effectuée à l'aide du kit NucleoSpin RNA (Macherey-Nagel). Une fois lysée, les cellules sont passées sur une colonne de filtration qui permet de réduire la viscosité de l'échantillon et d'éliminer les débris. Soixante dix pourcent d'éthanol sont ajoutés volume à volume de lysat cellulaire filtré. Le mélange est passé sur colonne de silice qui maintient les acides nucléiques après une étape de centrifugation pendant 30 secondes à 11000 g. Un lavage de la membrane est effectué pour éliminer les sels, suivi d'une première étape de digestion à la DNase pendant 15 minutes à température ambiante. La DNase est éliminée après plusieurs lavages puis une élution de la colonne est effectuée par ajout de 50µL d'eau RNase-free. Une deuxième digestion à la Dnase est effectuée à 37°C pendant 25 minutes pour dégrader les éventuels résidus d'ADN. Après inactivation de la Dnase, les échantillons sont passés sur des colonnes provenant du kit NucleoSpin RNA clean up XS (Macherey-Nagel) qui permet d'éluer dans de petits volumes d'eau et ainsi de concentrer les ARNs. Une PCR de recherche de contamination par l'ADNg ainsi qu'une qPCR pour quantifier l'expression de *SNCA* ont également été effectuées.

#### 3.3.1.3. Préparation de la librairie d'ADNg à séquençer

La préparation de la librairie correspond aux étapes de traitement des ARNs qui vont permettre de générer les fragments d'ADNg portant à leurs extrémités des adaptateurs qui vont permettre le séquençage. Les fragments séquencés constituent alors les reads qui sont alignés sur un génome de référence.

### 3.3.1.3.1. Préparation de la librairie pour le séquençage par le séquenceur SOLiD™ 4

La préparation de la librairie a été faite à l'aide du kit SOLiD Total RNA-Seq (Thermo-Fischer) en suivant les recommandations du fournisseur (figure 28).



**Figure 28: Préparation de la librairie pour le séquençage par le séquenceur SOLiD™ 4**

Après extraction, les ARNs totaux élués dans 30µL ont été séchés au Speed vac pour permettre leur concentration (paragraphe 3.3.1.2.1.). Cette étape est nécessaire à l'utilisation du kit RiboMinus™ Eukaryote Kit for RNA-Seq (Invitrogen) d'élimination des ARNr qui ne fonctionne que pour les très petits volumes. Sept microgrammes d'ARNs totaux concentrés ont été traités à l'aide du kit RiboMinus™ Eukaryote Kit for RNA-Seq pour enrichir en ARNm, les ARNr étant majoritaires dans la cellule et masquant les autres types de transcrits. Après élimination des ARNr à hauteur d'environ 95%, les ARNs restants (300ng en moyenne) ont été incubés 10 minutes à 37°C en présence de RNase III. Cette enzyme va cliver les ARNs et générer des fragments d'une taille allant de 35 nucléotides à quelques milliers de nucléotides selon les échantillons. Un passage sur colonne permet de purifier le mélange réactionnel et de récupérer les ARNs enrichis après élimination des ARNr. Cent nanogrammes d'ARNs enrichis sont ajoutés à un mélange de séquences d'oligonucléotides adaptatrices P1 et AI et d'une solution d'hybridation pendant 10 minutes à 37°C. Une ligase est ensuite ajoutée pour permettre la liaison des adaptateurs P1 et AI. Cette étape de liaison s'effectue à

16°C pendant 16 heures qui permettent de fixer les adaptateurs de part et d'autres des fragments d'ARNs. Ces derniers servent de matrice aux amorces oligonucléotides pour la réaction de transcription inverse mais également à la réaction d'amplification lors du séquençage. La synthèse des ADNc a été effectuée à l'aide de l'ARN polymérase ArrayScript™ Reverse Transcriptase (contenue dans le kit SOLiD Total RNA-Seq) pendant 30 minutes à 42°C. Par le biais du kit MinElute® PCR Purification Kit (Qiagen), les ADNc nouvellement synthétisés sont filtrés sur colonne afin d'éliminer le maximum d'éléments provenant des étapes antérieures et pouvant perturber le séquençage (notamment les nucléotides nécessaires à la synthèse d'ADNc complémentaires et les différents tampons). Les fragments d'ADNc ont ensuite été déposés sur gel de polyacrylamide Novex® pre-cast gel products (Invitrogen) qui permet de séparer les fragments d'ADNc selon leur taille. Le gel a ensuite été découpé pour récupérer les fragments dont la taille est comprise entre 150 et 250pb. Pour augmenter les chances de détection des ARNs faiblement exprimés dont les fragments d'ADNc sont en petite quantité, une PCR a été effectuée qui permet à la fois l'amplification des fragments d'ADNc et l'ajout de l'adaptateur P2 et d'un code barre spécifique à chaque échantillon. Ce code barre ou barcode est nécessaire pour déterminer la provenance des fragments d'ADNc.

#### 3.3.1.3.2. Préparation de la librairie pour le séquençage par le séquenceur Ion Proton™

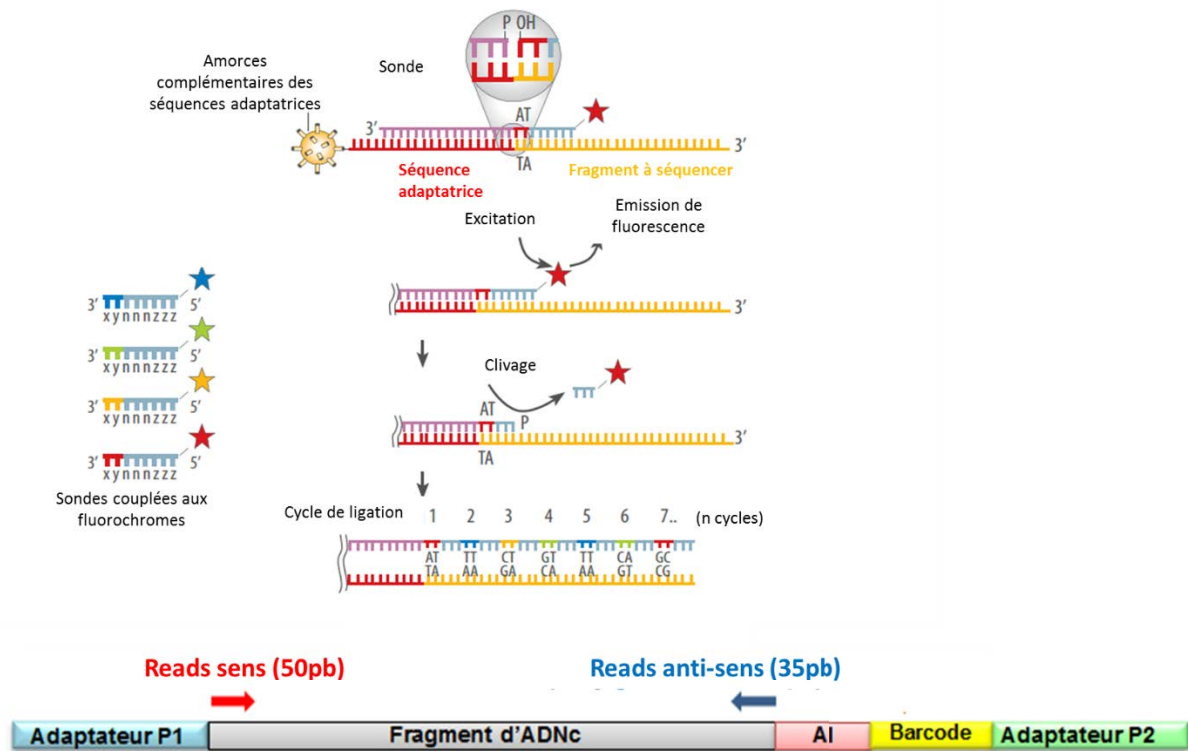
Le principe de la préparation de la librairie pour la procédure Ion Proton est le même que celui de la procédure SOLiD 4 à la différence que la quantité d'ARNs totaux de départ utilisée était de 5µg (et non pas 7µg). De plus, ce protocole ne requiert plus d'étape de sélection des fragments selon leur taille. Du fait de l'évolution des technologies, les kits ont évolué. Ainsi pour la préparation des librairies pour le séquençage Ion Proton, nous avons utilisé les kits Ion Total RNA-Seq Kit v2 (Thermo-Fisher) et rRNA using the RiboMinus™ Eukaryote System v2 (Thermo-Fisher) pour la déplétion des ARNr en suivant les recommandations du fournisseur et en ajoutant une seconde étape de déplétion pour réduire au mieux les ARNr.

#### 3.3.1.4. Réaction de séquençage

##### 3.3.1.4.1. Séquençage effectué par le séquenceur SOLiD™ 4

Le séquençage par le séquenceur SOLiD4 est un séquençage par liaison basée sur une amplification des fragments d'ADNc par PCR en émulsion associée à du pyroséquençage. La PCR en émulsion requiert la formation d'un mélange d'huile et de solution aqueuse qui va permettre la formation de réacteurs à l'intérieur desquels se trouvent des billes magnétiques couplées aux amorces oligonucléotides, l'ensemble des réactifs nécessaires à la réaction de PCR ainsi que les

fragments d'ADNc formés à l'issue de la préparation de la librairie. Ces amorces sont complémentaires des séquences adaptatrices P1 et P2 (rajoutées au cours de la préparation de la librairie). Elles vont permettre la capture des fragments d'ADNc contenus dans l'émulsion et permettre l'initiation de la réaction de PCR. Après appariement des fragments d'ADNc sur les billes, celles-ci sont récupérées puis déposées sur une plaque de verre où s'effectue le séquençage (figure 29).



**Figure 29: Principe du séquençage par le séquenceur SOLiD™ 4**

En 5', l'amplification se fait grâce à l'amorce oligonucléotide P1 couplée à la bille. Les cycles d'amplification vont permettre l'incorporation de sondes de 5 nucléotides (2 premiers nucléotides qui se fixent à l'ADNc à séquencer par complémentarité de base et 3 nucléotides universels qui se fixent indépendamment du principe de complémentarité). Chaque sonde est marquée par un fluorochrome, spécifique des deux premiers nucléotides de la sonde. Si ces 2 nucléotides se fixent à l'ADNc, le fluorochrome va être excité et émettre une fluorescence qui sera enregistrée par une caméra. Après lecture, une enzyme va cliver les 3 nucléotides universels et laisser sur l'ADNc les 2 bases complémentaires. A l'inverse, en absence de complémentarité, la sonde ne se fixe pas, il n'y a pas d'émission de fluorescence et un nouveau cycle de ligation va être initié avec une nouvelle sonde de séquence différente. La caméra va ainsi lire la fluorescence émise à chaque cycle et ainsi pouvoir déterminer l'enchaînement de nucléotides et donc la séquence de l'ADNc. La capacité du séquenceur

ne permet pas de séquencer le fragment d'ADNc d'un bout à l'autre de sa séquence. La séquence adaptatrice P2 va donc permettre la fixation de l'extrémité 3' du fragment d'ADNc sur la bille. Dans ce cas la réaction de PCR est initiée par une amorce oligonucléotide complémentaire de la séquence adaptatrice A1 qui est plus proche de l'extrémité du fragment d'ADNc à séquencer. Ainsi pour un même fragment d'ADNc généré à l'issue de la préparation de la librairie, on obtient 2 morceaux de séquences appelées reads pairés dont la taille est égale à 50 bases pour les reads à l'extrémité 5' (ce sont les reads sens) et 35 bases pour ceux à l'extrémité 3' (ce sont les reads anti-sens).

#### 3.3.1.4.2. Séquençage effectué par le séquenceur Ion Proton™

Le séquençage par le séquenceur Ion Proton est un séquençage par mesure ionique basée sur la détection de proton  $H^+$ . Les fragments d'ADNc à séquencer sont préalablement fixés à des sphères qui sont elles mêmes disposées sur des puces contenues dans le séquenceur. Ces puces contiennent 165 millions de puits remplis d'une solution au sein desquels les billes sont réparties et où va s'effectuer la réaction d'amplification par PCR en émulsion. La PCR va permettre l'incorporation de nucléotides complémentaires qui sont ajoutés les uns après les autres. A chaque incorporation, des ions  $H^+$  vont être libérés. Cette libération de protons va modifier le pH (potentiel hydrogène qui dépend de la concentration en  $H^+$  dans une solution) de la solution contenue dans le puit. Grâce à un capteur, le séquenceur va pouvoir mesurer le pH à chaque cycle d'amplification pour déterminer si le nucléotide nouvellement ajouté a été incorporé ou non. A la différence du SOLiD™ 4, la capacité du séquenceur Ion Proton lui permet de séquencer 200 bases à chaque extrémité du fragment d'ADNc.

#### 3.3.1.5. Analyse quantitative des défauts d'épissage

##### 3.3.1.5.1. A partir des données de séquençage SOLiD™ 4

L'analyse quantitative a été réalisée à l'aide du pack Cufflinks qui permet l'alignement des reads sur un génome de référence, la reconstitution des transcrits dont sont issus ces reads et de mesurer des niveaux d'expression. Ce pack contient un ensemble d'outil : TopHat, Cufflinks, Cuffcompare, Cuffmerge, Cuffquant et Cuffdiff, chacun ayant une fonction bien précise dans la caractérisation des séquences présentes dans un échantillon.

A partir de la séquence des reads obtenue lors du séquençage, le logiciel TopHat va procéder à l'alignement de ces séquences sur le génome de référence choisi. Ce génome contient l'ensemble de gènes et transcrits annotés qui vont permettre l'identification des gènes et transcrits présents dans un échantillon. En effet, les reads dont la séquence s'aligne sur le génome c'est-à-dire ceux qui trouvent une correspondance dans le génome de référence vont être identifiés comme provenant de



cette séquence dans le génome. Plus le nombre de reads alignés est grand, plus la couverture du génome est grande c'est-à-dire plus la proportion du génome qui peut être analysée est importante. A noter que les reads non alignés sont non utilisables. Le génome que nous avons choisi était le génome de University of California Santa Cruz (UCSC) human reference genome hg19 (RefSeq hg19). Une fois les séquences déterminées, Cufflinks va procéder à la reconstitution des transcrits en analysant les reads qui se chevauchent les uns les autres. Ces transcrits ainsi reconstitués vont être à nouveau comparés au génome de référence par Cuffcompare pour vérifier la qualité de la reconstitution. Pour pouvoir procéder à l'analyse d'expression différentielle, les informations concernant les transcrits reconstitués et provenant des différents échantillons sont regroupés grâce à l'outil Cuffmerge. Après un enregistrement des données par Cuffquant, l'outil Cuffdiff va procéder à l'analyse d'expression différentielle. Pour cela le logiciel détermine une valeur nommée FPKM (Fragments Per Kilobase of exon per Million fragments mapped) qui pour un transcrit donné correspond au nombre de reads alignés sur le transcrit par rapport à la longueur du transcrit. De cette manière, en comparant la valeur de FPKM entre différents échantillons, le logiciel effectue une mesure de l'expression différentielle de chaque transcrit. Pour déterminer le niveau d'expression d'un gène, Cuffdiff fait la somme de l'ensemble des FPKM des différents transcrits de ce gène. Un test exact de Fisher a été effectué pour identifier les gènes et les transcrits différentiellement exprimés de manière significative ( $p < 0,05$ ).

#### 3.3.1.5.2. A partir des données de séquençage Ion Proton™

Le pack Cufflinks a également été utilisé pour déterminer les gènes et transcrits dont l'expression diffère entre les malades et les témoins. La procédure est identique à celle décrite ci-dessus à la seule différence que le génome de référence qui a été utilisé ici est le génome Ensembl GRCh37.p13 (Genome Reference Consortium Human Reference 37).

#### 3.3.1.6. Analyse qualitative des défauts d'épissage

##### 3.3.1.6.1. A partir des données de séquençage SOLiD™ 4

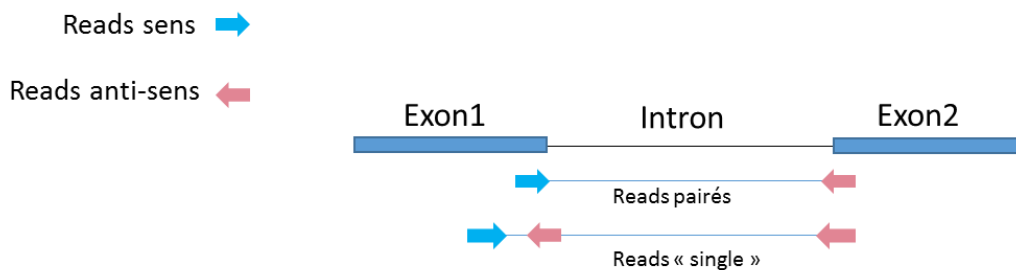
L'étude des jonctions entre exons et de leurs variations a été effectuée à l'aide du logiciel Bioscope™. A partir des données d'alignement obtenues grâce au logiciel TopHat, le logiciel Bioscope™ est capable d'identifier la position des reads sur le génome. Il fournit ainsi deux informations : une information quantitative sur le niveau d'expression des exons sur lesquels s'alignent les reads et une information qualitative par identification des jonctions entre exons. La mesure quantitative est définie par une valeur appelée RPKM (Reads Per Kilobase of exon model, per

Million mapped reads) correspondant au nombre de reads alignés sur un exon par rapport à la longueur de l'exon.

$$RPKM = 10^9 \times \frac{\text{Nombre de reads alignés sur l'exon}}{\text{Nombre de reads total} \times \text{Longueur de l'exon}}$$

Le logiciel permet également de déterminer les jonctions entre exons présentes dans un échantillon par comparaison des jonctions présentes dans le génome de référence. Pour qu'une jonction soit identifiée, 2 conditions doivent être remplies (figure 30) :

- Un read doit être aligné sur chacun des deux exons impliqués dans la jonction
- Un read doit être aligné à cheval sur la jonction entre ces deux exons.



**Figure 30 : Critères de détection des jonctions entre exons par le logiciel Bioscope™**

*A l'issue du séquençage, chaque fraquemnt d'ADNc à séquencer a généré 2 reads (un de chaque côté) ; on parle de reads pairés qui sont alignés sur le génome. Il existe des cas de figure où seul un des deux reads de la paire réussit à être aligné ; on parle alors de reads single. La longueur du fragment à séquencer étant relativement courte, l'alignement des reads pairés sur deux exons permet l'identification d'exons proches et l'alignement du reads single sur la jonction entre ces deux exons confirme qu'il s'agit d'exons successifs.*

Une jonction est définie comme « connue » si elle est présente dans le génome de référence ; à l'inverse la jonction est définie comme « potentiellement nouvelle » si cette dernière n'est pas retrouvée dans la base de données. Dans les deux cas, Bioscope™ attribue à la jonction identifiée une valeur appelée JCV (Junction Confidence Value) qui confère un degré de confiance à la jonction reconnue par Bioscope. Cette valeur va de 0 à 100, 100 représentant le plus haut degré de confiance attribuable à une jonction. Elle dépend de la couverture de séquençage et est déterminé grâce à l'algorithme suivant:

$$JCV_{j_{x-y}} = \sum_{i=1}^n PQV_i - 10 \log_{10} (EEM_{j_{x-y}})$$

Equation 1. Junction Confidence Value

$$EEM_{j_{x-y}} = \frac{RC_x}{\frac{\ell_x}{\mu_T + 3 \times \sigma_T}} \times \frac{RC_y}{\frac{\ell_y}{\mu_T + 3 \times \sigma_T}}$$

avec, PQV étant la qualité du séquençage ; RC étant le nombre de reads alignés sur l'exon ; l étant la longueur de l'exon,  $\mu$  étant la moyenne et T étant la déviation standard du fragment de séquence entre les 2 reads.

### 3.3.1.6.2. A partir des données de séquençage Ion Proton™

L'outil utilisé est l'outil Cuffcompare du pack Cufflinks. Après reconstitution des transcrits par l'outil Cufflinks, Cuffcompare compare les transcrits reconstitués aux transcrits présents dans le génome de référence. De cette manière, il identifie les similitudes et les différences entre ces deux groupes de transcrits (reconstitués *versus* présents dans le génome de référence) et leur attribue un code appelé « class code » (Tableau 7)

Class code	Description
=	Complete match of intron chain
c	Contained
j	Potentially novel isoform (fragment): at least one splice junction is shared with a reference transcript
e	Single exon transfrag overlapping a reference exon and at least 10 bp of a reference intron, indicating a possible pre-mRNA fragment
i	A transfrag falling entirely within a reference intron
o	Generic exonic overlap with a reference transcript
p	Possible polymerase run-on fragment (within 2Kbases of a reference transcript)
r	Repeat. Currently determined by looking at the soft-masked reference sequence and applied to transcripts where at least 50% of the bases are lower case
u	Unknown, intergenic transcript
x	Exonic overlap with reference on the opposite strand
s	An intron of the transfrag overlaps a reference intron on the opposite strand (likely due to read mapping errors)
.	.tracking file only, indicates multiple classifications

**Tableau 7: Différents types de transcrits identifiés après reconstitution par Cufflinks.**

Du fait de notre intérêt pour les variations de l'épissage, nous nous sommes particulièrement intéressés aux transcrits à class code j. Ils correspondent aux transcrits potentiellement nouveaux

ayant une composition en exons différentes des transcrits du même gène identifiés dans le génome de référence et pouvant ainsi résulter d'un défaut d'épissage.

### 3.3.1.6.3. A partir des données combinées de séquençage SOLiD™ 4 et Ion Proton™

L'étude de faisabilité et l'étude pilote ayant été effectuées à des temps différents, avec des outils et des génomes différents, nous avons choisi de faire une nouvelle analyse des variations qualitatives de l'épissage de façon simultanée avec le même outil. Pour cela nous avons calculé, pour chaque gène et chaque échantillon, le « score de défaut d'épissage ou Sd » comme définit par Bai *et al.*<sup>281</sup>. Ce score correspond à :

$$Sd = \frac{\frac{\text{Nombre de reads alignés sur les introns}}{\text{Longueur totale des introns}}}{\frac{\text{Nombre de reads sur les exons}}{\text{Longueur totale des exons}}}$$

Ce rapport est le reflet de l'incorporation et/ou de l'inclusion d'exons ou d'introns dans les transcrits des différents gènes exprimés. En comparant ce rapport entre différents échantillons, on peut avoir une information sur l'exclusion/inclusion des exons/introns et donc avoir une idée des épissages alternatifs présents dans l'échantillon.

### 3.3.1.7. Caractérisation des processus biologiques, des voies canoniques et des gènes dérégulés

#### 3.3.1.7.1. Logiciel Gene Set Enrichment Analysis (GSEA)

Les gènes significativement dérégulés ont été soumis au logiciel Gene Set Enrichment Analysis (GSEA, <http://www.broadinstitute.org/gsea/index.jsp>) qui permet de déterminer si les gènes dérégulés montrent une appartenance significative à une voie ou à un processus cellulaire donné. Pour ce faire, la liste de gènes dérégulés a été soumise au logiciel sous la nomenclature officielle HUGO. Le logiciel va ensuite interroger la base de données Gene Ontology pour associer un nom à un processus biologique donné. Il va ensuite calculer la statistique de recouvrement entre le nombre de gènes présents dans la liste entrée dans le logiciel et le nombre de gènes présents dans le processus cellulaire dans lequel il a été identifié.

#### 3.3.1.7.2. Logiciel Ingenuity Pathways Analysis (IPA®)

Les gènes au sein desquels de nouvelles jonctions putatives ont été identifiées par Bioscope™ ont été soumis au logiciel Ingenuity Pathway Analysis (IPA®, Ingenuity Systems, [www.ingenuity.com](http://www.ingenuity.com)). Ce logiciel permet d'identifier les voies canoniques perturbées par ces variations d'épissage qualitatives. Un test exact de Fisher calcule la valeur de probabilité  $p$  qui représente la probabilité de l'association entre les gènes de la liste soumise au logiciel et ceux de la voie canonique soit due au hasard. Une fois les voies canoniques identifiées, celles-ci ont été comparées aux voies canoniques identifiées à partir des listes de gènes dérégulés dans les données de puces d'expression de l'analyse du transcriptome des parkinsoniens porteurs de la mutation G2019S de *LRRK2* (étude chapitre 1), des sujets sporadiques (étude chapitre 1) et des parkinsoniens porteurs de mutations *ATXN2* (étude chapitre 2).

#### 3.3.1.7.3. Création de liste de gènes

Afin de compléter notre analyse de variation de l'épissage, nous avons généré une liste de gènes liés à l'épissage à partir de données provenant des données d'épissage SpliceFac (<http://rulai.cshl.edu/cgi-bin/SpliceFac/spliceFac.cgi?process=searchSFForm>) et SpliceAid ([http://193.206.120.249/splicing\\_tissue.html](http://193.206.120.249/splicing_tissue.html)). Ces informations ont été complétées avec des gènes liés à l'épissage provenant de la base de données Genecards et des données de la littérature de gènes de protéines à domaine de liaison aux ARNs<sup>282</sup>. Un gène pouvant contenir plusieurs noms cette liste contient à la fois les noms officiels des gènes (nomenclature HUGO) ainsi que leurs alias (autres appellations du même gène). A l'issue de ce regroupement, nous avons obtenu une liste de 4305 noms de gènes liés à l'épissage que nous avons croisée à nos listes de gènes et transcrits différentiellement exprimés ainsi qu'à nos listes de gènes ayant des modifications de jonctions entre exons dans le but de déterminer si les gènes de l'épissage étaient soumis à des variations de leur expression mais aussi à des perturbations de leurs propres événements d'épissage.

Nous avons également généré une liste de gènes d'intérêt dans la MP à partir des données de la littérature des gènes mutés dans la maladie ainsi que des facteurs de susceptibilité provenant de l'étude d'association de Nalls *et al.*<sup>217</sup>. Cette liste de 487 gènes a été croisée à la liste des gènes présentant des défauts d'épissage mis en évidence par le calcul du Sd afin d'identifier si les perturbations de l'épissage affectaient des gènes connus de la maladie.

### 3.3.1.8. Amplification par PCR, RT-PCR et RT-qPCR

#### Protocole de PCR:

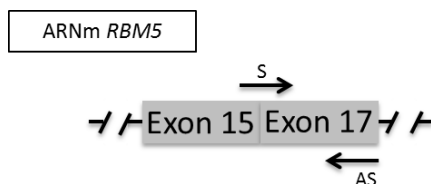
La recherche de contamination par l'ADNg des échantillons ARNs utilisés pour le séquençage a été effectuée en effectuant une PCR à partir de ces ARNs avec des amorces oligonucléotides introniques ciblant le gène *LRRK2*. Le protocole qui a été suivi est indiqué ci-dessous.

Étapes de PCR	Température	Durée	Nombre de cycles	
<b>Dénaturation</b>	95°C	10 minutes	1 cycle	
	95°C	1 minute		
<b>Élongation</b>	60°C	30 secondes	4 cycles	
	72°C	30 secondes		
	95°C	1 minute	20 cycles avec une diminution de 0,5°C à chaque cycle	
	60°C	30 secondes		
	72°C	30 secondes		
	95°C	1 minute	11 cycles	
	60°C	30 secondes		
	72°C	30 secondes		
		72°C	10 minutes	1 cycle
	<b>Terminaison</b>	4°C	10 minutes	1 cycle

La polymérase utilisée pour l'amplification étant une ADN polymérase ADN dépendante, l'amplification ne pouvait se faire qu'en présence d'ADNg dans les échantillons. La présence d'ADN a été vérifiée par dépôt sur gel d'agarose du produit de PCR qui en présence d'une contamination par l'ADNg devait révéler un amplicon de 332pb. Tous les échantillons que nous avons utilisés n'ont pas révélé cet amplicon et n'étaient donc pas contaminés par l'ADNg.

#### Protocole de RT-PCR:

Afin de valider l'existence des nouvelles jonctions identifiées par le logiciel Bioscope™, une RT-PCR a été effectuée avec des couples d'amorces d'oligonucléotides dont l'une des deux était à cheval sur la jonction nouvelle (lorsque la séquence au niveau des jonctions entre exons permettait la synthèse de ce type d'amorce, figure 31).



**Figure 31 : Hybridation d’amorces oligonucléotides de PCR sur la jonction nouvelle**

*La jonction entre l’exon 15 et l’exon 17 du gène RBM5 est une jonction nouvelle identifiée par le logiciel Bioscope™. Pour confirmer son existence, nous avons créé un couple d’amorces oligonucléotides avec au moins une des deux amorces (ici, l’amorce sens (S)) à cheval sur la jonction nouvelle. L’autre amorce (amorce anti-sens(AS)) se trouve à l’intérieur d’un exon. Pour les autres jonctions identifiées par le logiciel Bioscope que nous avons validées expérimentalement, la création des amorces oligonucléotides a été faite selon le même principe*

Les ADNc double brin utilisés pour la PCR ont été synthétisés par transcription inverse à l’aide du kit QuantiTect® Reverse Transcription (Qiagen) en suivant les recommandations du fournisseur. Les ARNs (soit 250ng) qui ont été utilisés sont ceux des 6 sujets étudiés au cours de l’étude de faisabilité (analyse RNAseq n°1). Pour l’amplification par PCR des fragments de séquence incluant la nouvelle jonction, une première étape de dénaturation des ADNc obtenus après transcription inverse a été effectuée par chauffage à 50°C pendant 2 minutes puis 95°C pendant 10 minutes. Dix cycles de « touch down » comprenant une étape à 95°C pendant 1 minute, une étape à 10 degrés au dessus de la température d’hybridation des amorces oligonucléotides pendant 30 secondes et 72°C pendant une minute ont permis d’augmenter la spécificité de l’amplification en favorisant l’appariement des amorces oligonucléotides au niveau de leur séquence cible. Trentes cycles d’amplification ont été effectués incluant une étape à 95°C pendant 1 minute, une étape à la température d’hybridation des amorces pendant 30 secondes et une étape à 72°C pendant une minute. Cette amplification a été achevée par une dernière étape à 72°C pendant 10 minutes.

Protocole de RT-qPCR :

Les variations quantitatives des transcrits ayant des nouvelles jonctions ont été validées sur l’appareil SpeOnePlus Real-Time PCR System (Applied Biosystems) en utilisant les mêmes amorces d’oligonucléotides que celles utilisées pour la RT-PCR (Tableau 8). La réaction de qPCR a été réalisée à partir de 0,4µL d’ADNc (synthétisés à partir de 250ng d’ARNs) dans un volume final de 10µL. Une étape de dénaturation du double brin d’ADNc a été effectuée par chauffage à 95°C pendant 10 minutes, permettant de rompre les liaisons hydrogènes entre les deux brins. Cette séparation a permis l’hybridation des amorces oligonucléotides complémentaires de chaque simple brin initiant ainsi la synthèse d’un nouveau ADNc double brin. Cette synthèse a été effectuée après 30 cycles de chauffage à 95°C pendant 15 secondes suivis d’une étape à la température d’hybridation des

amorces pendant 1 minute. Cette étape permet l'incorporation d'un marqueur fluorescent « le SYBR®Green (Applied Biosystems) » dont la fluorescence est théoriquement proportionnelle à la quantité d'ADNc au sein duquel le marqueur s'est inséré. La comparaison du signal de fluorescence entre différents échantillons permet de mettre en évidence des différences d'expression du gène ou de(s) transcrit(s) d'un gène quantifié(s). Le gène *MAN2B1* a été utilisé comme gène de référence, gène dont l'expression ne doit pas varier entre les différents échantillons, afin de normaliser les données d'expression. La méthode  $2^{-\Delta\Delta C_t}$ <sup>283</sup> a été utilisée pour effectuer le calcul d'expression différentielle. L'analyse statistique des rapports d'expression entre les individus malades et les témoins a été réalisée par un test de Mann et Whitney (seuil de significativité  $p < 0,05$ ).

Comme indiqué dans les paragraphes 3.3.1.2.1. et 3.3.1.2.2., une quantification de l'expression de *SNCA* a été effectuée après chaque extraction d'ARNs à partir des ARNs extraits. Le protocole suivi pour la RT-qPCR est le même que celui indiqué ci-dessus à la différence que les ARNs extraits et destinés à être séquencés par le séquenceur Ion Proton ont été convertis en ADNc à l'aide du kit High-Capacity cDNA Reverse Transcription (Life technologies) en suivant les recommandations du fournisseur.

Gènes	Séquence des amorces	Température d'hybridation (°C)	Taille de l'amplicon (pb)
<b>LRRK2 sens</b>	TTTTGATGCCTGACATAGTGGAC	60	322
<b>LRRK2 anti-sens</b>	CACATCTGAGGTCAGTGGTTATC		
<b>SNCA sens</b>	ACAGCAGTAGCCCAGAAGACAG	60	198
<b>SNCA anti-sens</b>	GGCTTCAGGTTTCGTAGTCTTGATA		
<b>RBM5 sens</b>	TGGCTCTCCGCAGTACC	62,6	100
<b>RBM5 anti-sens</b>	GCGAGTTGGGGTCATAATAG		
<b>CCLN1 sens</b>	CCCTTACAATGGCTAT	50	107
<b>CCNL1 anti-sens</b>	ATGATTATGATGTCTTCG		
<b>MAN2B1 sens</b>	GGAGAGGAGGCGGGATTA	61,4	123
<b>MAN2B1 anti-sens</b>	GGAGCGGTCAGTCAGCAC		

**Tableau 8: Amorces oligonucléotides utilisées pour les PCR, RT-PCR et RT-qPCR**



### 3.3.2. Résultats

#### 3.3.2.1. Étude de faisabilité (séquençage SOLiD™ 4)

##### 3.3.2.1.1. Analyse quantitative de l'expression des gènes

Dans le but de valider le potentiel de la technique de RNAseq dans l'identification de variations quantitatives et qualitatives de l'épissage, nous avons fait une première étude du transcriptome des cellules mononucléées périphériques sanguines de 3 parkinsoniens *versus* 3 témoins appariés en sexe et en âge. Le nombre de reads alignés sur le génome RefSeq hg19 est récapitulé dans le tableau ci-dessous (Tableau 9)

Sujets	3310121	3310167	3310199	P227/01	P59/04	P299/01
<b>Nombre total de reads alignés</b>	24 841 586	35 091 692	32 682 374	28 873 540	25 041 740	27 750 699

**Tableau 9: Nombre de reads générés par individu et alignés sur le génome humain RefSeq hg19**

*Le nombre de reads générés, par sujet, à l'issue de ce premier séquençage est répertorié dans le tableau ci-dessus. Les sujets 3310121, 3310167, 3310199 correspondent à des sujets témoins ; la patiente P227/01 correspond à une parkinsonienne porteuse de la mutation G2019S de LRRK2 et les malades P59/04 et P299/01 correspondent à deux patientes porteuses d'une duplication du locus de SNCA.*

Par utilisation du pack Cufflinks, nous avons mis en évidence des variations significatives de l'expression de 1438 gènes entre les malades et les témoins ( $p < 0,05$ ). Nous avons soumis cette liste de gènes au logiciel GSEA, afin d'identifier les processus dans lesquels ils étaient impliqués. Les processus biologiques identifiés sont récapitulés dans le tableau 10.

Processus biologiques	Nombre de gènes	FDR q-value
Transduction du signal	140	1,32E-29
Développement des organismes multicellulaires	92	1,17E-19
Développement des structures anatomiques	89	3,41E-19
Développement des systèmes	78	1,47E-17
Processus du système immunitaire	44	1,71E-15
Transduction du signal lié aux récepteurs de surface cellulaire	62	2,91E-15
Réponse au stress	54	5,21E-15
Réponse aux stimuli externes	41	1,87E-14
Régulation de la qualité biologique	47	5,63E-14
Signalisation cellule-cellule	46	5,86E-14
Prolifération cellulaire	52	9,15E-14
Développement des organes	55	1,13E-13
Réponse immunitaire	34	2,39E-13
Voie de signalisation des récepteurs couplés aux protéines G	41	2,68E-13
Régulation positive des processus biologiques	60	1,63E-12

**Tableau 10: Top 15 des processus biologiques liés aux gènes dérégulés (Analyse GSEA ; FDR<0,05)**  
*Pour chaque processus biologique identifié, le nombre de gènes qui s’y rattachent et dont l’expression varie entre les malades et les témoins a été indiqué.*

L’analyse de ces différents processus montre, que chez ces 3 malades, les gènes dont l’expression varie sont principalement associés à la transduction du signal, au développement et à la prolifération cellulaire par les processus tels que « Développement des organismes multicellulaires », « Développement des structures anatomiques » ou encore « Prolifération cellulaire », ainsi qu’à l’immunité par l’altération des « Processus du système immunitaire » et de la « Réponse immunitaire ». Ces processus rappellent des variations que nous avons déjà mises en évidence au cours de l’étude du transcriptome des sujets porteurs de la mutation G2019S de *LRRK2* mais aussi chez les sujets sporadiques.

Afin d’identifier si parmi ces gènes certains étaient en lien avec l’épissage, nous avons comparé la liste des 1438 gènes différentiellement exprimés entre les malades et les témoins à la liste des facteurs liés à l’épissage que nous avons générée. Cette comparaison a révélé une variation d’expression de 16 gènes liés à l’épissage entre les parkinsoniens et les témoins appariés dont 11 sont sous-exprimés soit 68,7% (Tableau 11).

Gènes	Ratio (M/T)	p-value
KIN	-2,15	5,00E-05
ZCCHC8	-1,68	9,50E-04
KCNJ5	-1,85	2,20E-03
PABPC1L2A	-2,74	2,30E-03
RBFOX1	-1,93	2,35E-03
RBM26	-1,55	3,95E-03
THOC1	-1,57	6,60E-03
RPL34-AS1	-1,76	9,85E-03
DAZL	2,37	1,28E-02
RBM19	-1,46	1,55E-02
CCNA1	1,56	2,15E-02
ELAVL2	1,38	3,84E-02
KPNA2	1,39	4,06E-02
TARDBP	-1,35	4,33E-02
TIA1	-1,35	4,45E-02
SETD1A	1,48	4,76E-02

**Tableau 11: Gènes de facteurs liés à l'épissage différentiellement exprimés entre les parkinsoniens et les témoins (Test de Fischer exact ;  $p < 0,05$ )**

*Ce tableau indique les 16 gènes liés à l'épissage et dont l'expression varie entre les malades et les témoins. Le ratio M/T représente le rapport des niveaux d'expression chez les malades (M) et les témoins (T). On observe une variation plus importante du gène PABPC1L2A, un analogue fonctionnel des protéines de la famille PABP.*

De façon intéressante, nous avons observé une diminution de l'expression du gène *RBFOX1* connu également sous le nom d'*A2BP1*, un facteur régulateur de l'épissage d'*ATXN2*. De plus, nous avons également observé la diminution de l'expression du gène *PABPC1L2A*, un gène appartenant à la famille des PABP protéines dont nous avons supposé l'intérêt dans la MP dans le chapitre 2. Cette famille ayant un rôle dans l'épissage mais aussi dans la traduction et la dégradation des ARNm par le NMD, la variation de l'expression de *PABPC1L2A* pourrait suggérer une perturbation de ces mécanismes chez les malades.

### 3.3.2.1.2. Analyse quantitative de l'expression des transcrits

De façon analogue à l'analyse de l'expression des gènes, nous avons observé la variation d'expression de transcrits correspondant à 1306 gènes impliqués dans les processus biologiques indiqués dans le tableau ci-dessous (Tableau 12).

Processus biologiques	Nombre de gènes	FDR $q$ -value
Transduction du signal	112	1,38E-19
Développement des structures anatomiques	78	3,64E-16
Développement des organismes multicellulaire	79	5,22E-16
Développement des systèmes	67	3,30E-14
Réponse au stress	48	4,12E-13
Transduction du signal liée aux récepteurs cellulaires de surface	54	7,91E-13
Signalisation cellule-cellule	39	6,19E-11
Développement des organes	47	7,93E-11
Régulation négative des processus biologiques	50	7,21E-10
Réponse au stimulus externe	32	1,07E-09
Voie de signalisation des protéines couplées aux protéines G	33	2,42E-09
Prolifération cellulaire	41	3,81E-09
Réponse à une agression	24	3,88E-09
Processus des systèmes immunitaires	32	3,88E-09
Régulation négative des processus cellulaires	46	9,86E-09

**Tableau 12: Top 15 des processus biologiques liés aux transcrits dérégulés (Analyse GSEA ; FDR<0,05)**

Cette analyse a montré la perturbation de la transduction du signal, de la prolifération cellulaire via les perturbations des processus biologiques reliés à la croissance et à la prolifération cellulaire tels que « Développement de structures anatomiques », « Développement des organismes multicellulaire », « Développement des organes », « Prolifération cellulaire », mais aussi des perturbations de la réponse au stress et de l'immunité. De façon intéressante, ces voies correspondent à celles altérées au niveau génique et correspondent également à des voies déjà mise en évidence dans nos précédentes analyses.

Parmi les transcrits dérégulés, 24 correspondent à des gènes de facteurs liés à l'épissage (Tableau 13). A noter que 14 de ces 24 gènes soit 58,3% correspondent à des gènes dont l'expression est elle-même perturbée au niveau génique (Tableau 13).

Gènes	Ratio (M/T)	p-value
<b>ZCCHC8</b>	-1,90	9,50E-04
<b>RBFox1</b>	1,87	2,15E-03
<b>KCNJ5</b>	11,59	2,20E-03
<b>PABPC1L2A</b>	1,38	2,30E-03
<b>RBM26</b>	1,92	3,95E-03
<b>RBM19</b>	-2,08	5,25E-03
<b>THOC1</b>	-1,66	6,60E-03
<b>RPL34-AS1</b>	-4,11	9,85E-03
CPSF1	-1,85	1,02E-02
HNRNPD	-3,22	3,12E-02
HNRNPA1L2	1,39	3,36E-02
ZRANB2	-2,74	3,57E-02
<b>ELAVL2</b>	-2,83	3,69E-02
DROSHA	-2,07	3,81E-02
SNRPB2	-1,55	3,89E-02
<b>KPNA2</b>	-1,76	4,06E-02
<b>TARDBP</b>	1,48	4,33E-02
SNRNP27	2,44	4,44E-02
CPSF7	2,55	4,46E-02
HNRNPF	-1,35	4,50E-02
DHX35	-1,57	4,61E-02
<b>SETD1A</b>	1,27	4,76E-02
GAR1	-1042,12	4,77E-02
<b>TIA1</b>	-166989,75	8,29E-01

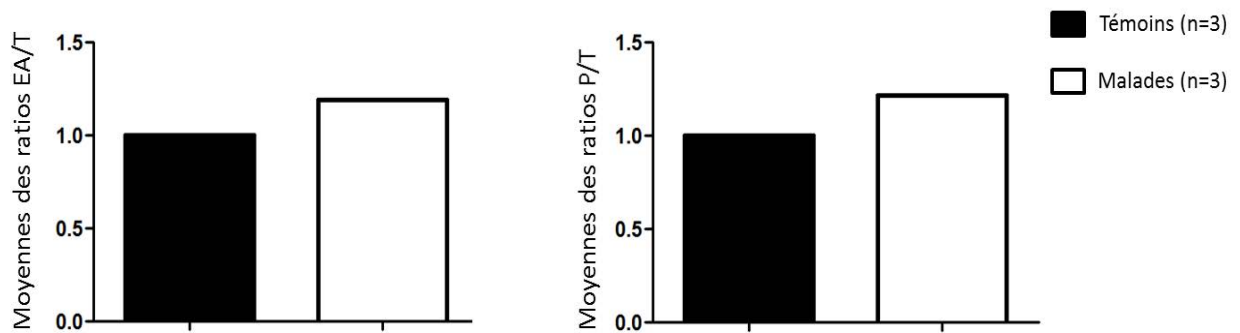
**Tableau 13: Transcrits de gènes liés à l'épissage différenciellement exprimés entre les malades et les témoins (Test de Fisher exact ;  $p < 0,05$ )**

*Ce tableau récapitule les 14 gènes liés à l'épissage ayant au moins un transcrit différenciellement exprimé entre les malades et les témoins. En rouge sont indiqués les gènes dont l'expression varie à la fois au niveau du gène et de ses transcrits.*

D'une manière générale, les variations quantitatives que nous avons observées dans cette analyse, reflètent des variations déjà observées dans nos études précédentes à savoir des perturbations de processus biologiques liés à la maladie comme la transduction du signal, la prolifération cellulaire et l'immunité ainsi que les perturbations du métabolisme des ARNs.

#### 3.3.2.1.3. Analyse qualitative de l'épissage : création de nouvelles jonctions entre exons

Une nouvelle fois, nos données ont révélé des variations quantitatives de l'expression de facteurs liés à l'épissage. Ceci nous a encouragés à poursuivre dans l'idée que des variations au niveau des jonctions entre exons existaient chez les parkinsoniens. Grâce aux données provenant de Bioscope™, nous avons pu déterminer le nombre de jonctions connues et le nombre de jonctions potentiellement nouvelles détectées dans chaque échantillon. A partir de ces données, nous avons calculé par échantillon, le rapport EA/T correspondant à l'ensemble des épissages alternatifs détectés dans un échantillon (jonctions connues et putatives, EA) par rapport à l'ensemble des jonctions présentes dans le génome de référence (T). Nous avons aussi comptabilisé le nombre de jonctions putatives (P) dans chaque échantillon et calculé le rapport P/T correspondant au rapport des jonctions putatives et de l'ensemble des jonctions présentes dans le génome. Ces rapports nous ont permis de faire une estimation des différences qu'il pouvait y avoir au niveau des jonctions entre exons entre les malades et les témoins. La moyenne des ratios EA/T des malades par rapport à celles des ratios des témoins montre une tendance à l'augmentation chez les malades (figure 32A). Ceci implique que le nombre d'événements d'épissage réalisés et détectés chez les malades est plus grand que chez les témoins. De façon similaire, la moyenne des ratios P/T des malades est sensiblement plus grande que celle des ratios des témoins suggérant que les variations qualitatives conduisant à la gènèse de nouvelles jonctions sont présentes en plus grand nombre chez les malades que chez les témoins (figure 32B).



**Figure 32 : Moyennes des ratios EA/T et P/T chez les témoins et chez les malades**

À gauche: Moyennes des ratios EA/T chez les témoins et les parkinsoniens ; à droite : Moyennes des ratios P/T chez les témoins et les parkinsoniens

Afin de valider la création de nouvelles jonctions entre exons chez les parkinsoniens, nous sommes intéressés à ces jonctions putatives. Bioscope™ fournit des informations sur les différentes jonctions entre exons présentes dans un échantillon auxquelles il attribue une JCV qui témoigne de la confiance que l'on peut accorder en l'existence de la jonction identifiée. Aucun critère n'étant définis pour déterminer quel seuil de JCV (confiance accordée à la jonction identifiée) choisir pour avoir le maximum de confiance en l'existence de ces jonctions, nous avons choisi de considérer l'ensemble des jonctions et de valider l'existence de jonctions d'intérêt par RT-PCR.

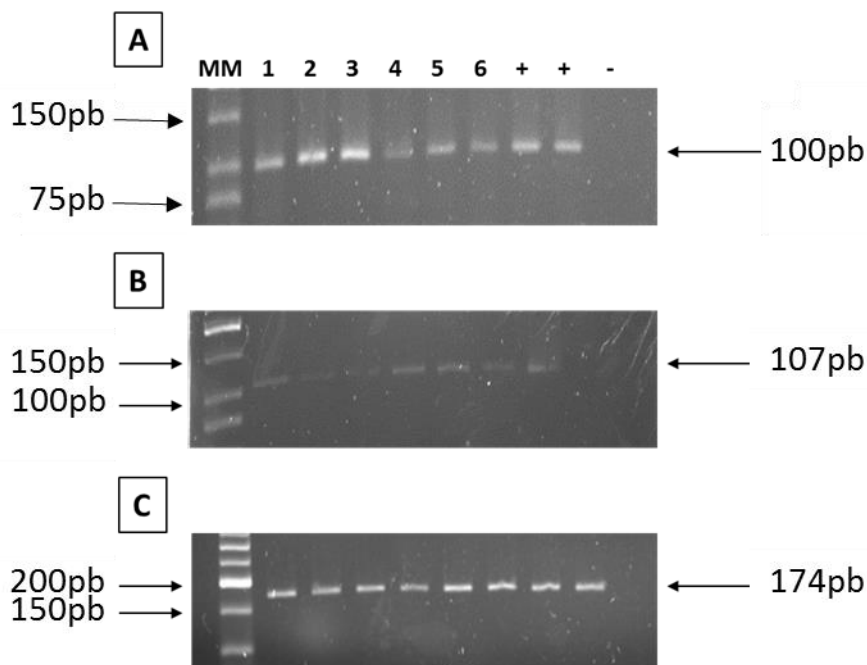
Trois mille deux cent seize gènes à jonctions nouvelles ont été identifiés (sans sélection sur la JCV) et croisés avec la liste des 4305 facteurs liés à l'épissage que nous avons générée. Ces jonctions nouvelles sont retrouvées dans 181 gènes de la liste des gènes liés à l'épissage. Nous avons à nouveau retrouvé des variations au niveau de gènes auxquels nous nous étions intéressés auparavant comme le gène *ATXN2*, *PABPC1*, et *PABPC4*. Ayant observé des variations d'expression du gène et de transcrit d'*A2BP1* (facteur jouant sur l'épissage d'*ATXN2*), il est possible que ces variations soient à l'origine des modifications de jonctions entre exons de l'*ATXN2* bien que cela mérite d'être vérifié.

Afin de confirmer l'existence de ces jonctions nouvelles, nous en avons sélectionnées certaines d'entre elles à différentes JCV (Tableau 14) que nous avons validées par RT-PCR lorsque la synthèse d'amorces à cheval sur la nouvelle jonction a été possible (Figure 33)

Gènes	Jonctions nouvelles	Sujets possédant la nouvelle jonction	JCV
<b>RBM5</b>	Exons 15-17	P299/01	35,21
<b>CCLN1</b>	Exons 9-11	P227/01 et P59/04	100 et 100
<b>RPS9</b>	Exons 3-5	P227/01	100

**Tableau 14: Jonctions nouvelles validées par RT-PCR**

*Ces jonctions nouvelles correspondent à des épissages de type cassette-exon (paragraphe 1.1.5.1.) avec exclusion de l'exon 16 de RBM5, 10 de CCLN1 et 4 de RPS9. D'après les données Bioscope ces nouvelles jonctions semblent n'être présentes que chez certains sujets car la nouvelle jonction de RBM5 a été détectée uniquement chez la patiente P59/04, celle de CCLN1 uniquement chez les patientes P227/01 et P59/04 et celle de RPS9 n'a été détectée que chez la patiente P227/01.*



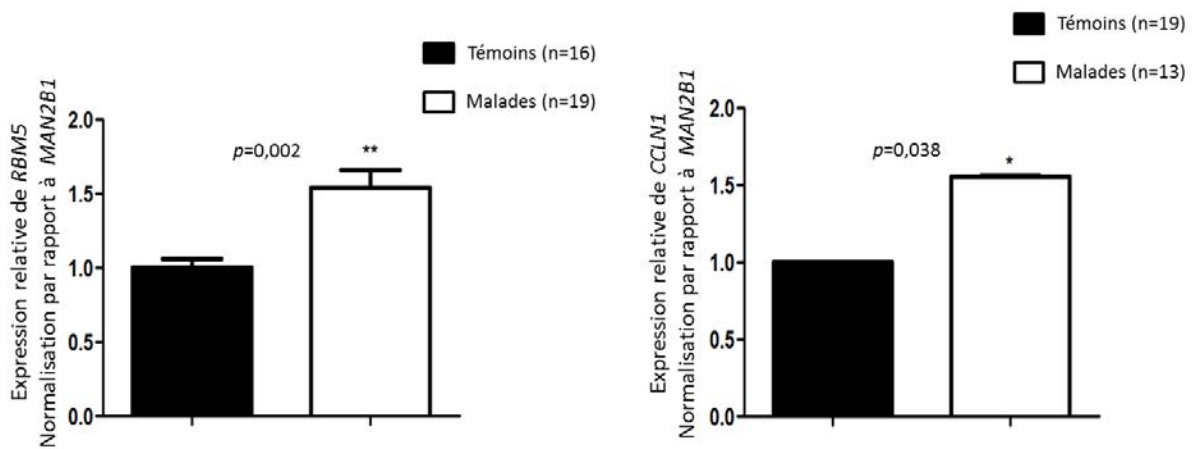
**Figure 33 : Amplicons de RT-PCR après dépôt sur gel d'agarose**

*Les fragments de séquence contenant la nouvelle jonction identifiée dans le(s) transcrit(s) de RBM5 (A), CCLN1 (B) et RPS9 (C) ont été obtenus après amplification par PCR des ADNc provenant des 3 malades et des 3 témoins de l'analyse de RNAseq n°1. 1 : 3310199 ; 2 : P299/01 ; 3 : P59/04 ; 4 : P227/01 ; 5 : 3310167 ; 6 : 3310199 ; + : Témoin positif ; - : Témoin négatif (H<sub>2</sub>O) ; MM : marqueur de poids moléculaire*

Les résultats de RT-PCR confirment l'existence de ces nouvelles jonctions qui sont étonnement retrouvées chez l'ensemble des sujets. Cependant, les différences d'intensité du signal sur le gel d'agarose suggèrent une différence d'expression de(s) transcrit(s) possédant ces nouvelles jonctions ; ceci sous-réserve que le dépôt sur gel eu été uniforme. Pour tester cette hypothèse, nous avons



effectué des RT-qPCR qui ont révélé une augmentation significative de(s) transcrit(s) de *RBMS* possédant la nouvelle jonction ainsi que pour ceux de *CCLN1* (figure 34).



**Figure 34 : Quantification par RT-qPCR des transcrits de *RBMS* et *CCLN1***

Ces quantifications ont été effectuées sur les ADNc provenant d'un groupe de sujets sélectionnés pour effectuer les validations. \* $p < 0,05$  ; \*\* $p < 0,01$  ; Mann et Whitney

La quantification par RT-qPCR de *RPS9* n'a pas été faite car les amorces sur la jonction nouvelle ne remplissaient pas les critères d'utilisation en qPCR (courbe de fusion à pics multiples), ce d'autant que le gel d'agarose n'a pas montré de variations d'intensité notables (figure 31C). Ces résultats suggèrent que les jonctions nouvelles de *RBMS* et *CCLN1* ne sont pas des jonctions nouvellement créées chez les parkinsoniens du fait d'un défaut de l'épissage puisqu'elles sont également présentes chez les témoins. Cependant les différences observées en RT-qPCR indiquent qu'il existe des variations de l'épissage de ces 2 gènes qui, chez les malades, accentuent la formation de ces jonctions par rapport à ce qui est observé chez les témoins. Il est toujours possible que, pour d'autres gènes, les variations qualitatives de l'épissage soit un phénomène de tout ou rien c'est-à-dire conduisant à la création de jonction n'existant que chez un des deux groupes (malades ou témoins).

Afin de voir si les changements au niveau des jonctions entre exons peuvent être responsables de perturbations conduisant à la maladie, nous avons analysé les voies canoniques dans lesquelles étaient impliqués ces gènes à nouvelles jonctions. La liste des 3216 gènes à jonctions nouvelles a été soumise à IPA® ; les voies canoniques identifiées ont été comparées aux voies canoniques révélées par nos analyses antérieures (Tableau 15).

Voies canoniques	Nouvelles jonctions <i>p</i> -value	Sporadiques <i>p</i> -value	G2019S <i>LRRK2</i> <i>p</i> -value	<i>ATXN2</i> <i>p</i> -value
<b>Voie de signalisation EIF2</b>	7,94E-13	3,72E-04	2,00E-15	
Voie d'ubiquitination des protéines	1,58E-10	4,68E-02	3,39E-04	
<b>Voie de régulation d'eIF4 et p70S6K</b>	2,40E-10	4,79E-02	6,31E-07	4,07E-02
Voie Rac	7,76E-09	3,39E-02	2,57E-02	
<b>Voie de signalisation PI3K/AKT</b>	3,54E-06	3,80E-02	4,47E-03	1,26E-02
Activation RAR	3,80E-06	1,62E-02	2,82E-03	
Super-voie des composés inositol-phosphates	3,89E-06	2,04E-02	1,51E-02	
Voie de signalisation CTLA4 dans les lymphocytes T cytotoxiques	4,07E-06	1,26E-02	1,70E-02	
Voie du cancer de la prostate	8,13E-06	2,51E-03	1,00E-02	
Biosynthèse des 3-phosphoinositides	1,29E-05	7,76E-03	4,37E-02	
Voie de signalisation des 14-3-3	4,68E-05	1,15E-02	1,41E-02	4,79E-02
Voie de signalisation de la leucémie myéloïde chronique	1,07E-04	1,78E-02	2,51E-02	
Signalisation des céramides	2,09E-03	6,92E-03	4,68E-02	
Signalisation des acides docosahexaénoïques	2,75E-03	2,40E-02	4,79E-02	3,39E-02
Signalisation AMPK	3,80E-03	8,32E-05	3,89E-02	
Troubles Mitochondriaux	1,48E-02	1,51E-02	1,86E-07	

**Tableau 15: Voies canoniques communes aux différentes analyses du transcriptome (Analyse IPA ;  $p < 0,05$ )**

Ces nouvelles jonctions sont retrouvées au sein de gènes représentant des voies de signalisation connues dans la physiopathologie de la maladie tels que la dégradation par le système ubiquitine-protéasome « Voie d'ubiquitination des protéines », les voies de survie cellulaire « Voie Rac », « Voie de signalisation PI3K/Akt », « Voie de signalisation des 14-3-3 », « Signalisation AMPK », l'immunité « Voie de signalisation CTLA4 dans les lymphocytes T cytotoxiques » ou encore la mitochondrie. Une fois de plus, nous avons retrouvé des voies communes à celles de nos précédentes analyses et notamment les voies de « régulation de eIF4 et p70S6K » et « signalisation PI3K/Akt » commune à l'ensemble des formes étudiées jusqu'à présent ainsi que la voie de « signalisation EIF2 » qui était la voie commune aux sujets sporadiques et aux parkinsoniens porteurs de la mutation G2019S de *LRRK2* la plus fortement dérégulée.

### 3.3.2.2. Étude pilote (séquençage Ion Proton™)

#### 3.3.2.2.1. Analyse quantitative de l'expression des gènes

Afin de voir si les perturbations de l'épissage étaient observées dès les premiers stades de la maladie, nous avons fait une nouvelle analyse à partir d'échantillon provenant de sujets à des stades précoces. Six malades et 5 témoins appariés en sexe et en âge ont été séquencés, cette fois-ci avec le séquenceur Ion Proton. Les reads ont été alignés sur le génome Ensembl GRCh37 et leur nombre par échantillon est indiqué dans le tableau ci-dessous (tableau 16).

**A**

Sujets	3110149	3110138	3110150	3110147	3110025
Nombre de reads	67 401 669	74 339 599	74 339 599	65 826 035	76 005 259

**B**

Sujets	3310250	3310268	3310260	3310245	3310242	3310239
Nombre de reads	87 127 167	57 428 119	70 703 395	66 932 786	78 715 727	77 051 227

**Tableau 16: Nombre de reads par sujet, alignés sur le génome humain Ensembl GRCh37**

*A : Parkinsoniens sporadiques (stade précoce); B : Sujets contrôles*

L'analyse des données a été effectuée selon le même procédé que celui employé lors de l'étude de faisabilité. Nous avons ainsi observé la variation d'expression de 437 gènes ( $p < 0.05$ ) associés aux processus cellulaires indiqués ci-dessous (Tableau 17).

<b>Processus biologiques</b>	<b>FDR q-value</b>
Transduction du signal	1,81E-10
Développement des organismes multicellulaires	4,57E-10
Régulation des processus métaboliques cellulaires	8,65E-06
Régulation des processus métaboliques	8,81E-06
Régulation négative des processus biologiques	3,78E-05
Régulation des processus métaboliques des acides nucléiques et nucléobase nucleoside nucleotide	4,10E-05
Régulation négative des processus cellulaires	6,25E-05
Régulation de l'expression génique	9,86E-05
Transcription	1,14E-04
Processus métaboliques des acides nucléiques et nucléobase nucléoside	1,64E-04
Transcription de promoteurs à ARN polymérase II	2,32E-04
Cascade de signalisation intracellulaire	2,66E-04
Régulation de la transcription ADN dépendante	9,62E-04
Processus métaboliques des biopolymères	1,01E-03
<b>Processus de biosynthèse des ARNs</b>	<b>1,55E-03</b>
Transcription ADN dépendante	1,55E-03
<b>Processus du métabolisme des ARNs</b>	<b>2,75E-03</b>
Établissement de la localisation	3,82E-03

**Tableau 17: Processus biologiques liés aux gènes dérégulés chez les parkinsoniens à stades précoces (Analyse GSEA ; FDR<0,05)**

De façon intéressante, nous avons retrouvé des perturbations de gènes liés au métabolisme de ARNs qui suggèrent que des altérations de ces mécanismes surviennent dès les premiers stades de la maladie.

Nous avons donc croisé la liste de gènes différentiellement exprimés entre les malades et les témoins à la liste des 4305 annotations de gènes liés à l'épissage. Les variations quantitatives de l'épissage chez ces patients aux premiers stades de la maladie impliquent les variations d'expression de douze gènes (Tableau 18).

Gènes	Description	Ratio (M/T)	p-value
<b>RBMS3</b>	RNA Binding Motif, Single Stranded Interacting Protein 3	-4,78	2,75E-03
<b>PPIL3</b>	Peptidylprolyl Isomerase (Cyclophilin)- Like 3	-2,12	3,00E-04
<b>KHDRBS2</b>	KH Domain Containing, RNA Binding, Signal Transduction Associated 2	-2,01	3,81E-02
<b>RPL9</b>	Ribosomal Protein L9	1,33	4,49E-02
<b>CCDC94</b>	Coiled-Coil Domain Containing 94	1,36	3,74E-02
<b>RPL41</b>	Ribosomal Protein L41	1,37	4,24E-02
<b>SF3B4</b>	Splicing Factor 3b, Subunit 4, 49kDa	1,44	3,90E-02
<b>RPL28</b>	Ribosomal Protein L28	1,70	4,60E-03
<b>RPL23AP7</b>	Ribosomal Protein L23a Pseudogene 7	1,86	2,95E-02
<b>RBFOX3</b>	RNA Binding Protein, Fox-1 Homolog (C. Elegans) 3	1,96	2,65E-02
<b>LGALS2</b>	Lectin, Galactoside-Binding, Soluble, 2	1,52	9,60E-03
<b>U3</b>	U3 Small Nucleolar Ribonucleoprotein	-1,63	2,20E-03

**Tableau 18: Douzes facteurs liés à l'épissage dérégulés chez les parkinsoniens (Test de Fisher exact ;  $p < 0,05$ ).**

*Le ratio M/T représente le rapport des niveaux d'expression du gène considéré dans le groupe des malades par rapport à celui du groupe des témoins*

Ces gènes diffèrent de ceux mis en évidence par l'étude pilote. Toutefois, il est important de noter la variation d'expression d'*U3*, un constituant majeur de la machinerie d'épissage. Il est donc possible que des perturbations au sein même de la machinerie effectrice de l'épissage soit à l'origine d'un grand nombre de modifications de ce processus dans la cellule.

#### 3.3.2.2.2. Analyse quantitative de l'expression des transcrits

Les transcrits différentiellement exprimés ont été analysés et correspondent à 357 gènes ( $p < 0,05$ ). Comme pour les gènes, le nombre de transcrits identifiés est nettement plus faible que ceux identifiés lors de l'étude de faisabilité (1438 gènes et 1306 gènes à transcrits différentiellement exprimés dans l'étude de faisabilité contre 437 gènes dérégulés et 357 gènes à transcrits différentiellement exprimés dans l'étude pilote, paragraphe 3.3.2.2.1.). Contrairement aux précédentes analyses où le premier processus identifié était la « transduction du signal », l'analyse des processus biologiques des transcrits différentiellement exprimés chez les sujets à des stades

précoces de la maladie, a montré que le premier processus biologique atteint était le système immunitaire (Tableau 19).

Processus biologiques	Nombre de gènes	FDR q-value
Réponse immunitaire	12	1,09E-05
Développement des organismes multicellulaires	23	1,96E-05
Processus du système immunitaire	13	2,07E-05
Régulation négative des processus cellulaires	16	2,44E-04
Prolifération cellulaire	14	2,97E-04
Régulation négative des processus biologiques	16	2,97E-04
Transduction du signal	26	3,35E-04
Développement des structures anatomiques	19	7,16E-04
Cascade de signalisation intracellulaire	15	7,45E-04
Régulation de la prolifération cellulaire	10	1,04E-03
Régulation des processus métaboliques de la cellule	16	1,06E-03
Régulation des processus métaboliques	16	1,12E-03
Processus de biosynthèse cellulaire	10	1,12E-03
Réponse à l'hypoxie	4	1,12E-03
Développement cellulaire	13	1,76E-03
Processus métaboliques des biopolymères	24	2,12E-03
Régulation de la transcription des promoteurs à ARN polymérase II	9	2,30E-03
Transcription de promoteurs à ARN polymerase II	11	3,33E-03

**Tableau 19: Top 18 des processus biologiques liés aux transcrits dérégulés (Analyse GSEA ; FDR<0,05)**

Par ailleurs, les perturbations de la prolifération cellulaire, de la transduction du signal et de la réponse au stress identifiés précédemment, sont également retrouvées à des stades précoces de la maladie (Tableau 17).

Le croisement de la liste des transcrits dérégulés avec la liste des gènes liés à l'épissage montre que 12 gènes liés à l'épissage sont dérégulés entre les malades et les témoins (Tableau 20). Ces transcrits diffèrent également de ceux identifiés par la première analyse.

Gènes	Description	Ratio (M/T)	p-value
<b>CCDC94</b>	Coiled-Coil Domain Containing 94	1,37	3,67E-02
<b>EIF4G1</b>	Eukaryotic Translation Initiation Factor 4 Gamma, 1	N.D	7,10E-03
<b>KHDRBS2</b>	KH Domain Containing, RNA Binding, Signal Transduction Associated 2	-2,24	4,70E-02
<b>RPL23AP7</b>	Ribosomal Protein L23a Pseudogene 7	2,42	1,20E-02
<b>RPL28</b>	Ribosomal Protein L28	1,81	4,04E-02
<b>RPL28</b>	Ribosomal Protein L28	2,25	4,24E-02
<b>RPL9</b>	Ribosomal Protein L9	2,01	2,68E-02
<b>RPLP0</b>	Ribosomal Protein, Large, P0	N.D	4,47E-02
<b>RPS11</b>	Ribosomal Protein S11	N.D	7,55E-03
<b>VIM</b>	Vimentin	N.D	4,40E-03
<b>YBX1</b>	Y Box Binding Protein 1	1,39	1,90E-02
<b>DDX42</b>	DEAD (Asp-Glu-Ala-Asp) Box Helicase 42	N.D	3,72E-02
<b>U3</b>	Small nucleolar RNA U3	8,06	4,46E-02
<b>U3</b>	Small nucleolar RNA U3	-1,62	2,20E-03

**Tableau 20: Transcrits de gènes liés à l'épissage (Test de Fisher exact ;  $p < 0,05$ )**

*N.D (non déterminé): les ratios d'expression n'ont pas pu être calculés car la valeur de FPKM dans l'un des deux groupes (malades ou témoins) est égale à 0*

Bien que l'analyse des processus biologiques n'ait pas fait ressortir de perturbations en lien avec l'épissage, certains transcrits de gènes liés à ce processus sont différentiellement exprimés entre les malades et les témoins. Parmi ces transcrits, nous avons observé la variation d'expression de deux transcrits du gène *U3* que l'analyse au niveau génique avait déjà révélé. Les variations des gènes et transcrits liés à l'épissage dans l'analyse de faisabilité nous ont révélé que ces variations quantitatives étaient associées à des changements de compositions en exons des transcrits. Nous avons donc testé si des modifications similaires pouvaient être observées dès les premiers stades de la maladie.

### 3.3.2.2.3. Analyse qualitative de l'épissage : création de nouvelles jonctions entre exons

De même que précédemment, nous nous sommes interrogés sur les conséquences des variations quantitatives de gènes et transcrits liés à l'épissage sur la création de nouvelle jonction. Après analyse des données fournies par cuffcompare (3.3.1.6.2.), nous avons observé la création de 28 transcrits à class code J dérégulés entre les malades et les témoins ( $p < 0,05$ ), soit 28 transcrits potentiellement nouveaux. A noter que la majorité de ces transcrits potentiellement nouveaux correspondent à des gènes impliqués dans l'immunité (Tableau 21) qui semble être perturbée dès les premiers stades de la maladie (Tableau 17).

Gènes	Description	Ratio (M/T)	p-value
<b>AL161772.1</b>	N.A	1,93	1,50E-04
<b>B3GNT7</b>	UDP-GlcNAc:BetaGal Beta-1,3-N-Acetylglucosaminyltransferase 7	-1,86	4,50E-04
<b>IGKV1-12</b>	Immunoglobulin Kappa Variable 1-12	2,463	6,00E-04
<b>FCRL2</b>	Fc Receptor-Like 2	2,17	2,75E-03
<b>VIM</b>	Vimentin	N.D	4,40E-03
<b>APOBEC3A</b>	Apolipoprotein B mRNA Editing Enzyme, Catalytic Polypeptide-Like 3A	-1,58	6,20E-03
<b>CLC</b>	Charcot-Leyden Crystal Galectin	1,83	6,75E-03
<b>LGALS2</b>	Lectin, Galactoside-Binding, Soluble, 2	-1,52	9,60E-03
<b>RBKS</b>	Ribokinase	-2,19	9,70E-03
<b>IGKV1-13</b>	Immunoglobulin Kappa Variable 1-13 (Gene/Pseudogene)	2,49	1,11E-02
<b>IGLC7</b>	Immunoglobulin Lambda Constant 7	1,80	1,98E-02
<b>RP11-106J23.1</b>	N.A	6,50	2,41E-02
<b>PDPR</b>	Pyruvate Dehydrogenase Phosphatase Regulatory Subunit	1,41	2,57E-02
<b>CRIP2</b>	Cysteine-Rich Protein 2	2,52	2,69E-02
<b>UTP20</b>	UTP20, Small Subunit (SSU) Processome Component, Homolog (Yeast)	1,74	2,76E-02
<b>RN7SL197P</b>	RNA, 7SL, Cytoplasmic 197, Pseudogene	-1,47	3,00E-02
<b>DUS1L</b>	Dihydrouridine Synthase 1-Like (S. Cerevisiae)	23,48	4,08E-02
<b>CCNB1IP1</b>	Cyclin B1 Interacting Protein 1, E3 Ubiquitin Protein	N.D	4,14E-02



Ligase			
<b>IGKV1-6</b>	Immunoglobulin Kappa Variable 1-6	1,73	4,19E-02
<b>BHLHE40</b>	Basic Helix-Loop-Helix Family, Member E40	-1,54	4,21E-02
<b>GAB2</b>	GRB2-Associated Binding Protein 2	-60,18	4,34E-02
<b>IGHV3-52</b>	Immunoglobulin Heavy Variable 3-52 (Pseudogene)	-3,67	4,48E-02
<b>TBC1D27</b>	TBC1 Domain Family, Member 27	1,69	4,58E-02
<b>FCRL5</b>	Fc Receptor-Like 5	2,42	4,77E-02
<b>FTSJ3</b>	FtsJ Homolog 3 (E. Coli)	N.D	4,78E-02
<b>DCAF11</b>	DDB1 And CUL4 Associated Factor 11	N.D	4,88E-02

**Tableau 21: Vingt-huit transcrits à jonctions nouvelles, dérégulés entre les malades et les témoins (Test de Fisher exact ;  $p < 0,05$ )**

*N.A (non assigné) : aucune description n'est fournie pour ce gène dans NCBI. N.D (non déterminé): les ratios d'expression n'ont pas pu être calculés car la valeur de FPKM dans l'un des deux groupes (malades ou témoins) est égale à 0*

### 3.3.2.3. Étude des données combinées de séquençage SOLiD™ 4 et Ion Proton™

Les différences observées entre les résultats des deux analyses précédentes peuvent s'expliquer en partie par le fait que les sujets utilisés étaient différents mais aussi par le fait que les méthodes de séquençage, le génome de référence et les outils bio-informatiques d'analyse d'expression différentielle et de jonctions entre exons diffèrent. Nous avons donc refait une analyse en mesurant le score de déficience d'épissage « Sd » pour l'ensemble des sujets utilisés au cours des deux précédentes analyses auxquels ont été greffé des sujets supplémentaires (paragraphe 3.3.1.1.). Au total, cette analyse correspond à une comparaison de 19 malades *versus* 15 témoins. Comme indiqué paragraphe 3.3.1.6.3., le calcul du Sd a été déterminé à partir des données de séquençage de chaque individu qui certes proviennent de séquenceurs différents mais qui a le mérite d'analyser les variations qualitatives de l'épissage selon la même méthode et sur l'ensemble des échantillons diminuant ainsi les biais de l'analyse. De plus, le nombre de sujets utilisés pour cette étude étant plus grand ceci devrait augmenter la robustesse et la fiabilité des variations observées et augmenter ainsi les chances que ces défauts d'épissage soient retrouvés dans d'autres échantillons.

#### 3.3.2.3.1. Analyse qualitative de l'épissage : création de nouvelles jonctions entre exons

Cent quarante-cinq gènes ont révélé un Sd significatif (t-test ;  $p < 0,05$ ). Comme pour les précédentes analyses, cette liste a été soumise à GSEA pour identifier les processus biologiques affectés par ces défauts qualitatifs de l'épissage (Tableau 22).

Processus biologiques	Nombre de gènes	FDR <i>q</i> -value
Processus métabolique des protéines cellulaires	16	5,93E-5
Processus métaboliques des macromolécules cellulaires	16	5,93E-5
Processus métaboliques des protéines	16	1,22E-4
Traduction	7	1,83E-4
Processus métaboliques des biopolymères	18	2,23E-4
Processus biosynthétiques cellulaires	8	5,54E-4
Transcription	11	1,25E-3
Processus métaboliques des acides nucléiques et nucléobase nucleoside nucleotide	14	1,29E-3
Processus biosynthétiques des macromolécules	7	3,53E-3
Processus biosynthétique	8	5,05E-3
Modifications post-traductionnelles des protéines	8	5,05E-3
Régulation des processus métaboliques	10	5,93E-3
Régulation des processus métaboliques	10	6,2E-3
Réponse au stress	8	6,22E-3
Régulation de l'expression génique	9	7,39E-3
Processus du métabolisme des ARNs	10	7,63E-3
Régulation de la transcription	8	1,05E-2
Apoptose	7	1,05E-2
Mort cellulaire programmée	7	1,05E-2
Processus cellulaires cataboliques	5	1,47E-2

**Tableau 22: Top 20 des processus biologiques liés aux gènes à défauts d'épissage (Analyse GSEA ; FDR<0,05)**

Les perturbations de l'épissage affectent des gènes impliqués dans de nombreux processus relatifs aux métabolismes des ARNs tels que la traduction, la transcription mais aussi mécanismes de réponse aux stress et à la mort cellulaire.

Afin de voir si certains de ces gènes sont impliqués dans l'épissage, la liste des gènes à Sd significatif a été croisée avec la liste des gènes liés à l'épissage que nous avons contruite (tableau 23).

Gènes	Description	Valeur du t-test
<b>PQBP1</b>	Polyglutamine Binding Protein 1	7,11E-03
<b>C14orf166</b>	Chromosome 14 Open Reading Frame 166	8,31E-03
<b>RPL39</b>	Ribosomal Protein L39	1,00E-02
<b>SNRPC</b>	Small Nuclear Ribonucleoprotein Polypeptide C	1,17E-02
<b>POLR2A</b>	Polymerase (RNA) II (DNA Directed) Polypeptide A	1,30E-02
<b>RPS17</b>	Ribosomal Protein S17	1,32E-02
<b>RPS9</b>	Ribosomal Protein S9	1,61E-02
<b>LSM10</b>	LSM10, U7 Small Nuclear RNA Associated	2,25E-02
<b>RBM22</b>	RNA Binding Motif Protein 22	2,44E-02
<b>SNRPG</b>	Small Nuclear Ribonucleoprotein Polypeptide G	2,54E-02
<b>NELFE</b>	Negative Elongation Factor Complex Member E	2,58E-02
<b>XAB2</b>	XPA Binding Protein 2	2,83E-02
<b>RPL27</b>	Ribosomal Protein L27	3,08E-02
<b>RPS21</b>	Ribosomal Protein S21	3,44E-02
<b>PUF60</b>	Poly-U Binding Splicing Factor 60KDa	3,45E-02
<b>RPL3</b>	Ribosomal Protein L3	3,52E-02
<b>WTAP</b>	Wilms Tumor 1 Associated Protein	4,00E-02
<b>RPL34</b>	Ribosomal Protein L34	4,09E-02
<b>IMP4</b>	IMP4, U3 Small Nucleolar Ribonucleoprotein	4,18E-02

**Tableau 23: Gènes de facteurs liés à l'épissage ayant des perturbations de leur épissage (t-test ; p<0,05)**

Bien que les gènes impactés soient une nouvelle fois différente des précédentes analyses, l'ensemble de nos différentes études montre qu'il existe des défauts d'épissage entre les malades et les témoins qui affectent de nombreux gènes mais également des gènes effecteurs de l'épissage.

Afin de savoir si les variations qualitatives impactent non pas des voies mais des gènes connus pour avoir un rôle dans la maladie, nous avons à nouveau croisé la liste des gènes ayant un Sd significatif avec la liste que nous avons générée contenant 437 gènes porteurs de mutations ou ayant un polymorphisme associé à la MP (Tableau 24).

Gènes	Description	Valeur du t-test
<b>BAK1</b>	BCL2-Antagonist/Killer 1	3,73E-03
<b>BST1</b>	Bone Marrow Stromal Cell Antigen 1	4,92E-03
<b>LGALS2</b>	Lectin, Galactoside-Binding, Soluble, 2	1,40E-02
<b>C16orf13</b>	Chromosome 16 Open Reading Frame 13	1,54E-02
<b>BTN3A2</b>	Butyrophilin, Subfamily 3, Member A2	2,86E-02
<b>RPL3</b>	Ribosomal Protein L3	3,52E-02
<b>PTPN18</b>	Protein Tyrosine Phosphatase, Non-Receptor Type 18 (Brain-Derived)	3,88E-02
<b>BCL7B</b>	B-Cell CLL/Lymphoma 7B	3,92E-02
<b>RGS18</b>	Regulator Of G-Protein Signaling 18	4,46E-02
<b>HSD17B11</b>	Hydroxysteroid (17-Beta) Dehydrogenase 11	4,49E-02
<b>IGSF6</b>	Immunoglobulin Superfamily, Member 6	4,73E-02
<b>LOC101927814</b>		4,73E-02
<b>METTL9</b>	Methyltransferase Like 9	4,73E-02

Tableau 24: Gènes de la maladie de Parkinson avec un Sd significatif (t-test ;  $p < 0,05$ )

### 3.3.3. Discussion

Les résultats que nous avons obtenus au cours du chapitre 1 et du chapitre 2 nous avaient suggéré l'existence de défauts d'épissage dans la MP. Dans la présente étude, nous avons en effet identifié des variations d'expression de gènes et de transcrits liés à l'épissage entre les malades et les témoins. Ces changements d'expression sont présents à la fois chez des sujets présentant des formes génétiques de la maladie et chez des sujets sporadiques corroborant notre hypothèse selon laquelle les variations liées à l'épissage impacteraient un grand nombre de cas de la maladie. De plus, nos données montrent que ces perturbations de l'épissage n'impliquent pas uniquement des variations quantitatives mais concernent également de variations qualitatives par modifications de la composition en exons des transcrits. L'équipe de Shehadeh et ses collaborateurs ont également montré que les variations d'expression des facteurs d'épissage et en particulier du facteur SRRM2 étaient associées à des modifications qualitatives de l'épissage dans les cellules sanguines de parkinsoniens. Toutefois, nous n'avons pas retrouvé ces changements dans nos données. De façon intéressante, l'ensemble de ces modifications de l'épissage ont été observées à la fois chez des sujets à des stades précoces et des stades plus avancés de la maladie montrant que les perturbations de l'épissage sont présentes dès les stades précoces. Scherzer et ses collaborateurs avaient déjà suggéré que des modifications de l'épissage dans les cellules mononucléées périphériques sanguines de parkinsoniens avaient lieu chez les malades dès les premiers stades<sup>240</sup>. Toutefois, leur étude avait été menée à partir d'échantillons de 55 patients à un stade Hoehn et Yahr moyen égal à 2,3 ; stade au cours duquel les manifestations motrices sont d'ores et déjà visibles et ce de manière bilatérale. Ils ne témoignent donc pas des stades précoces de la maladie. En revanche, bien que nos effectifs soient moindres, les sujets aux stades précoces que nous avons étudiés ont un score Hoehn et Yahr moyen inférieur ou égal à 1,5 reflétant ainsi les perturbations moléculaires précoces. L'analyse d'Alieva *et al.* des processus biologiques liés aux gènes différentiellement exprimés dans les cellules sanguines de 4 patients *de novo* sans traitements et 4 témoins appariés a montré une association de 514 gènes avec l'épissage alternatif<sup>284</sup> corroborant l'existence de perturbations de ce processus dès les premiers stades.

Bien que nos données soient préliminaires et aient été obtenues à partir de petits effectifs, elles nous ont permis de mettre en évidence de nombreuses variations à la fois quantitatives et qualitatives. Les données quantitatives d'expression génique ont révélé des variations de gènes et de transcrits impliqués dans la transduction du signal, la réponse immunitaire, la prolifération et la croissance cellulaire ainsi que la réponse au stress. Ces voies avaient déjà été observées dans nos données de puces de sujets porteurs de la mutation G2019S de *LRRK2* et de cas sporadiques tendant

ainsi à valider l'existence de ces dérèglements dans les cellules sanguines de parkinsoniens. Des modifications d'expression génique liée à la transmission synaptique et à la réponse immunitaire ont aussi été observées dans le cerveau de parkinsoniens porteurs de la même mutation de *LRRK2* ainsi que des sujets sporadiques<sup>242,227,279</sup>. L'ensemble de ces données souligne l'importance du système immunitaire dans la MP sachant que les gènes du système HLA font partie des gènes dérégulés dans le cerveau des parkinsoniens et que des polymorphismes de ces gènes sont associés à un risque accru de développer la maladie<sup>285,242</sup>. Il semblerait que des perturbations du système immunitaire soient présentes dès les stades précoces puisque nous avons observé des variations de gènes et transcrits impliqués dans ces processus chez les sujets *de novo*. Des variations d'expression de gènes liés à la réponse immunitaire ont été retrouvées par une autre équipe ayant étudié le transcriptome de cellules sanguines de 4 patients *de novo* versus les 4 témoins appariés suggérant l'implication du système immunitaire dès les premiers stades. Ces dérèglements de l'immunité ont également été retrouvés dans les cellules sanguines des sujets asymptomatiques porteurs de la mutation G2019S de *LRRK2*<sup>244</sup> et dans la substance noire de cas incidents à des stades de Braak entre 1 et 2 (lorsque la substance noire n'est pas encore atteinte) mimant ainsi les premiers stades de la maladie. Bien qu'étant menées sur de petits échantillons, les résultats de ces études convergent vers l'implication de perturbations du système immunitaire dès les premiers stades de la maladie. En nous intéressant à ces changements qualitatifs de l'épissage, nous avons observé que ces modifications de jonctions entre exons impactaient des voies et des processus connus pour être liés à la maladie, comme la « voie d'ubiquitination des protéines », les troubles mitochondriaux et l'immunité avec des perturbations de la « voie CTLA4 dans les lymphocytes T cytotoxiques ». Tout comme les variations quantitatives de gènes et de transcrits, ces jonctions nouvelles sont également observées dès les premiers stades et modifient également des jonctions entre exons de gènes impliqués dans la réponse immunitaire comme les gènes des immunoglobulines *IGKV1-12*, *IGKV1-13*, *IGKV1-6*, *IGHV3-52* ou des récepteurs aux immunoglobulines *FCRL2* et *FCRL5*.

Au-delà des voies et processus classiquement énoncés comme étant liés à la maladie, les données de RNAseq ont montré des perturbations de voies précédemment identifiées dans nos données de puces d'expression à savoir des perturbations de la « voie de signalisation EIF2 », de la voie de « régulation de eIF4 et p70S6K » ou encore de la « voie PI3K/Akt ». Ceci a permis de révéler que ces voies étaient également modulées par des perturbations de l'épissage des gènes qu'elles contiennent suggérant que les possibles perturbations de la traduction et de l'épissage dans la MP puissent être la résultante des perturbations de ces voies. De façon intéressante, nos croisements des listes de gènes ayant des perturbations qualitatives de l'épissage avec la liste de gènes liés à la physiopathologie de la maladie (gènes à mutations causales ou avec des polymorphismes) ont

montré que ces défauts d'épissage impactent aussi des gènes liés à la maladie. Ce lien entre l'épissage et la MP est renforcé par l'observation de nouveaux épissages au sein des gènes *ATXN2* et *PABP* (ayant un lien probable avec la MP) qui méritent d'être étudiés afin de voir si des modifications de l'épissage de ces gènes sont pathogènes. L'hypothèse selon laquelle les mutations causales ou les facteurs de risque de la MP puissent être associés à des variations de l'épissage est d'autant plus pertinente que des polymorphismes dans la région 3' non traduite de l'ARNm de *SNCA* sont responsables d'une augmentation du ratio isoformes 112/140 dans le cerveau de parkinsoniens sporadiques<sup>286</sup>, de même que dans les lignées cellulaires dopaminergiques de neuroblastome SH-SY5Y et un modèle murin de la MP exposés à du MPTP<sup>275</sup>. Des délétions au niveau des sites d'épissage des gènes *PINK1*<sup>287</sup> et *DJ-1*<sup>288</sup> sont responsables de modifications de la structure des transcrits exprimés. D'autres gènes tels que *LRRK2* et *PRKN* montrent une expression préférentielle de certains transcrits chez les parkinsoniens argumentant l'importance de l'épissage dans les processus physiopathologiques.

L'un des objectifs de ces études de RNAseq étaient de mettre en évidence les différents paramètres à prendre en compte pour l'emploi de cette méthode dans l'étude de l'épissage. L'un des facteurs pouvant influencer les résultats de RNAseq est le nombre de reads générés et alignés sur le génome de référence qui va définir la couverture de séquençage. La couverture correspond au nombre de gènes détectés par le RNAseq par rapport au nombre de gènes présents dans le génome de l'organisme étudié ; plus ce nombre est grand et plus la proportion de gènes et transcrits étudiés est importante, reflétant ainsi la réalité du transcriptome étudié. Des stratégies ont été développées pour évaluer la couverture qui serait suffisante pour obtenir des données de RNAseq fiables<sup>289</sup>. Dans notre cas, bien que le nombre de reads aient été supérieur dans l'étude pilote par rapport à l'analyse de faisabilité, la quantité de gènes et/ou de transcrits identifiés s'est avérée nettement inférieure par rapport à celle identifiée dans l'étude de faisabilité. Le changement du génome RefSeq pour le génome Ensembl pourrait être une explication à cette baisse. En effet, plusieurs bases des données d'annotations du génome humain sont actuellement disponibles incluant AceView Genes, Ensembl Genes, H-InvDB Genes, RefSeq Genes, UCSC Known Genes et Vega Genes. La principale différence entre ces bases de données est la quantité d'informations qu'elles contiennent. La base de données RefSeq semble être celle qui contient le moins d'informations mais qui présente l'avantage par rapport aux autres d'être une base de données non redondante ce qui facilite l'alignement des reads et donc l'identification des séquences auxquelles ces derniers correspondent. Po-Yen Wu *et al.* ont en effet montré, en comparant des données de séquençage de mêmes échantillons alignés sur des génomes différents, que l'annotation des reads était facilitée par la réduction de la complexité du génome de référence<sup>290</sup>. En effet, en contenant un nombre moins important de séquence, ceci réduit

la possibilité d'alignement des reads à plusieurs endroits dans le génome et donc la difficulté à déterminer la séquence dont ces reads sont issus. Nous avons en effet pu voir que le choix du génome de référence était déterminant puisque même si nos analyses sont effectuées sur des sujets différents nous nous attendions à obtenir quelques gènes communs ce qui n'a pas été le cas. En replaçant ceci dans le contexte d'une identification de biomarqueurs, il pourrait alors s'avérer difficile d'identifier une ou plusieurs cibles potentielles dont les variations pourraient être vérifiées chez différents groupes de sujets. Nous avons d'ailleurs rencontré ce problème puisque les gènes *RBM5* et *CCLN1* pour lesquels nous avons validé la surexpression de transcrits possédant de nouvelles jonctions, n'a pas été retrouvée dans nos différentes analyses. La nécessité de recourir à une méthode standardisée est plus qu'indispensable à la vue de notre objectif qui vise à maximiser les chances de découvrir des cibles potentielles fiables.



### 3.4. Chapitre 4 : Rôle potentiel de la traduction dans la maladie de Parkinson

Au cours de l'analyse de nos différents résultats, il nous est apparu que des modifications de la traduction puissent être impliquées dans la MP. En effet, que ce soit par la variation d'expression de gènes impliqués dans la voie de signalisation EIF2, par l'implication potentiel des protéines PABP ou encore par les variations qualitatives de l'épissage de gènes impliqués dans la synthèse protéique, nos données s'accordent à dire que des perturbations de la traduction seraient impliquées dans la physiopathologie de la MP.

Nous avons donc décidé de faire une revue des données de la littérature corroborant l'hypothèse de l'existence de perturbations de la traduction dans la MP. Nos recherches bibliographiques nous ont amené à la conclusion que les perturbations de la traduction associées à la maladie impliqueraient des gènes connus de la MP tels que *LRRK2*, *PRKN*, *PINK1*, *DJ-1*, *VPS35*, *ATXN2*, ou encore *EIF4G1* bien que son rôle dans la physiopathologie ait été revu depuis la publication de cette revue. Ces gènes contrôlèrent la synthèse protéique en modulant le recrutement de facteurs d'initiation de la traduction ou permettraient à la cellule de modifier son protéome en réponse à un stress cellulaire.

La traduction est un processus cellulaire hautement régulé faisant intervenir une machinerie cellulaire que sont les ribosomes et un grand nombre de facteurs protéiques associés. L'un des modes de régulation mis en place par la cellule est la phosphorylation des facteurs de traduction qui permet de moduler leur recrutement au niveau du site de synthèse protéique. Grâce à son activité kinase, l'orthologue de *LRRK2* chez la drosophile a été montré comme étant capable d'activer la traduction en inhibant l'interaction du répresseur 4E'BP avec la machinerie de traduction<sup>139</sup>. Ce lien entre *LRRK2* et la traduction a également été observé dans des lignées cellulaires humaines où cette protéine interagit avec plusieurs acteurs de la traduction dont la protéine ribosomique Rps15<sup>291</sup>. D'autres protéines codées par des gènes importants dans la maladie semblent jouer un rôle dans la traduction. C'est le cas de la protéine ataxine-2 qui, au-delà des résultats obtenus dans notre étude présentée dans le chapitre 2, est connue pour sa capacité à interagir avec les ribosomes et avec la protéine PABPC1 qui favorise la circularisation de l'ARNm à traduire et facilite l'exécution de la traduction. Les gènes porteurs de mutations récessives tels que *PRKN*, *PINK1* et *DJ-1* codent également des protéines pouvant jouer un rôle dans la synthèse protéique. Alors que *DJ-1* a été montré comme étant associée à des protéines ribosomiques et influençant l'efficacité de traduction chez *E. Coli*<sup>292</sup>, les protéines mitochondriales *PINK-1* et parkin modulent la traduction en réponse à un stress oxydant<sup>293</sup>.

Le contrôle de la traduction passe aussi par l'intermédiaire de modifications épigénétiques qui influent sur la nature du protéome telles que la méthylation et les miARNs. Les miARNs sont bien connus pour interagir avec leurs ARNm cibles et bloquer leur traduction en empêchant le recrutement de la machinerie de traduction et en induisant la dégradation des ARNm cibles. LRRK2 et ataxine-2 sont capables d'interagir avec des protéines essentielles à la reconnaissance des ARNm par les miARNs<sup>294,295,161</sup>; cette interaction laisse supposer que le rôle dans la traduction de certains des gènes de la MP pourrait également passer par une action au niveau de l'ARNm et pas uniquement au niveau des facteurs de traduction. De plus, la traduction de l'ARNm de *SNCA* a été rapportée comme étant différente entre les sains et les malades atteints de MP chez qui une hypométhylation de cet ARNm a été associée à une surexpression de l'ARN et de la protéine correspondante<sup>93</sup>. Ceci est d'autant plus intéressant que la méthyltransférase DNMT1, responsable de la méthylation de l'ARNm de *SNCA*, est séquestré dans le cytoplasme où elle ne peut agir sur ses cibles<sup>94</sup>. Si la méthylation renforce l'importance de la traduction dans la MP, les modifications affectant l'extrémité 3'UTR des ARNm accentue également le rôle de ce processus dans la maladie. En effet, Rhinn *et al.* ont montré que la longueur de l'extrémité 3'UTR des transcrits de *SNCA* jouait sur leur localisation cellulaire et leur capacité à s'accumuler<sup>296</sup>; les transcrits à longue extrémité 3' étant plus prompt à s'aggréger. Ces données sont cohérentes avec le fait que la surexpression de *SNCA* soit pathologique et laissent suggérer que certains transcrits seraient plus fortement traduits que d'autres, donnant ainsi à la traduction un rôle clé dans la physiopathologie de la MP.

**[Article 3: Deregulation of protein translation control, a potential game-changing hypothesis for Parkinson's disease pathogenesis](#)**

## **Deregulation of protein translation control, a potential game-changing hypothesis for Parkinson's disease pathogenesis**

Jean-Marc Taymans, Aurore Nkiliza and Marie-Christine Chartier-Harlin

UMR-S1172 Jean-Pierre Aubert Research Center – (INSERM - CHRU de Lille - Université de Lille), Early Stages of Parkinson's Disease Team, Lille, France.

*Corresponding authors:* Jean-Marc Taymans ([jean-marc.taymans@inserm.fr](mailto:jean-marc.taymans@inserm.fr)) and Marie-Christine Chartier-Harlin ([marie-christine.chartier-harlin@inserm.fr](mailto:marie-christine.chartier-harlin@inserm.fr))

### **Abstract**

Protein translation is one of the most fundamental and exquisitely controlled processes in biology, and is energetically demanding. The deregulation of this process is deleterious to cells, as evidenced by several diseases caused by mutations in protein translation machinery. Emerging evidence now points to a role for protein translation in the pathogenesis of Parkinson's disease (PD), a debilitating neurodegenerative movement disorder. In this paper, we propose a hypothesis that protein translation machinery, PD associated proteins and PD pathology are connected in a functional network linking cell survival to protein translation control. This hypothesis is a potential game changer in the field of the molecular pathogenesis of PD, with implications for the development of PD diagnostics and disease-modifying therapies.

**Keywords:** translation initiation factor (eIF), translation elongation factor (eEF), ribosome, EIF4G1, LRRK2, alpha-synuclein.

### *Molecular pathogenesis of Parkinson's disease*

Research on molecular disease pathogenesis is a complex endeavor involving the identification of disease factors in patients then investigating in depth how these factors affect cellular functions. By confirming the relevance of affected processes in disease models it is ultimately possible to exploit these findings to develop novel diagnostics and therapeutics for patients. In the past 15 years, a wealth of genetic studies has identified several disease factors for Parkinson's disease (PD), a debilitating and incurable neurodegenerative disease. Scientists in the PD field agree that the cause of neurodegeneration lies in a chronic deregulation of cellular processes causing accumulations of subtle cytotoxic effects ultimately leading to premature dopaminergic neuron death. Several cellular processes have been extensively studied in this regard, the most notable of which are protein folding, mitochondrial physiology, membrane physiology, vesicular transport, gene transcription, protein degradation or autophagy. Here, we discuss the hypothesis that protein translation problems are an integral part of Parkinson's disease pathogenesis and that further detailed investigation of protein translation in PD is warranted.

### *Protein translation*

Protein translation, one of the most fundamental and exquisitely controlled processes in biology, has received little attention in PD. Protein translation is the overall process whereby proteins are synthesized from the genetic information of RNA transcripts (Figure 1). This process is divided into 3 phases: initiation, elongation and termination. Initiation involves the concerted action of multiple translation initiation factors (eIFs; see glossary). When loaded with guanosine triphosphate (GTP), eIF2 recruits tRNAs to the 40S ribosomal subunit. eIF3 and eIF4 initiation factors are also recruited to the 40S and this complex of 40S with tRNA and eIFs, called the pre-initiation complex, can then bind to the 5' cap structure of mRNAs. The pre-initiation complex scans the 5' untranslated region (UTR) to find the start codon, at which point the 60S ribosomal subunit is recruited and initiation factors are released. It should be noted that there is also a cap-independent translation initiation mechanism where ribosomes are not recruited to the 5' cap but rather to a region of the mRNA presenting a sequence capable of recruiting ribosomes and/or translation initiation factors such as an internal ribosomal entry site (IRES) [1] or cap-independent translation element (CITE) [2]. During the elongation phase, elongation factors are recruited to the ribosome and a sequence of events is repeated for the addition of each new amino acid to the polypeptide chain: positioning the appropriate tRNA to the ribosome acceptor site (A-site), forming the new peptide bond and shifting the ribosome complex by 1 codon on the mRNA chain. The elongation process is subject to ribosome pausing, for instance in the case of co-translational folding of the nascent protein. When the translation complex reaches the stop codon, elongation stops and the ribosome is released from the

mRNA with the help of release factors. While the majority of mRNAs are translated in the cytosol or across endoplasmic reticulum membranes, some are also translated in mitochondria or on mitochondrial surfaces [3].

Several inherited diseases are caused by mutations in genes involved in the protein translation process. For instance, mutations in the ribosomal protein RPS19 are known to be a cause of Diamond-Blackfan Anemia [4]. Mutated protein translation proteins are also the cause of brain disorders, such as autism spectrum disorder [5], neurodegenerative disorders such as Charcot-Marie-Tooth disease (where mutations in tRNA synthetases are reported) [6] or for forms of cerebellar ataxia (where mutations are reported in translation initiation factor eIF2B) [7]. Other evidence for the role for protein translation in ataxia comes from the ataxin-2 gene (*ATXN2*) that is linked to spinocerebellar ataxia 2 (SCA2). Ataxin-2 protein binds ribosomes [8] as well as the PolyA Binding Protein Cytoplasmic 1 (PABPC1) protein [9] that interacts with eukaryotic translation initiation factor 4G (eIF4G), thereby favoring transcript circularization and translation initiation [10]. Ataxin-2 is also reported to interact with GW182 and Argonaute 1 (Ago1), two proteins involved in the biogenesis of microRNAs [11], and is also located in cellular compartments where protein translation occurs, such as P-bodies [12] and rough endoplasmic reticulum [13]. Interestingly, some mutations in *ATXN2* are associated with a PD phenotype, suggesting that disturbances in protein translation may contribute to PD pathogenesis.

#### *Genetic links between PD and protein translation*

Although the etiology of PD is not fully understood, genetic studies in patients have identified several mendelian and susceptibility genes involved in disease. A first series of disease genes were identified as causative factors of rare monogenic forms of PD. These include autosomal dominant genes such as the gene encoding the alpha-synuclein protein (*SNCA*), leucine-rich repeat kinase 2 (*LRRK2*) and *EIF4G1* or recessive genes such as the parkin gene (*PRKN*), Pten-induced kinase 1 (*PINK1*) or *DJ-1*. More recently, genome wide association studies, (GWAS) have identified genomic loci where single nucleotide polymorphisms are associated with increased risk for PD. Meta analysis has confirmed at least 22 candidate risk loci [14], including at the *SNCA*, microtubule associated protein tau (*MAPT*), *LRRK2* and glucosidase beta acid (*GBA*) genes.

Recent reports have shown that several proteins implicated in PD intervene in protein translation processes. The most obvious example is *EIF4G1*, which linked to both PD [15] and Lewy body dementia [16]. *EIF4G1* is a translation initiation factor, and is known to play a role as a scaffold in the eIF4F translation initiation complex that recruits ribosomes and tRNAs to the 5' cap structure of

mRNA [17]. Several mutations in EIF4G1 have been identified in PD patients, including p.A502V and p.R1205H [15], however these mutations are also present in healthy individuals, suggesting incomplete penetrance of these mutations. Testing of the functional effects of these pathogenic variants on protein translation has begun, and the eIF4G1 p.R1205H mutation deregulates protein translation (Chartier-Harlin, M.-C., unpublished and [18]). More recently, eIF4G1 was found to interact genetically in yeast with another PD gene, the vacuolar protein sorting-associated protein 35 (*VPS35*), indicating that protein deregulation pathways of eIF4G1 involves other genetic factors of PD [19].

Further links between PD and translation come from studies investigating the LRRK2 signaling pathway. LRRK2 can phosphorylate several proteins of the protein translation machinery, such as translation repressor protein 4E-BP in *Drosophila* [23] and ribosomal protein s15 in mammalian cell lines [21]. LRRK2 is also reported to interact with protein translation proteins, such as translation initiation factors (such as eIF2C1, eIF2C2, eIF4B, eIF3J) [20] and elongation factors (eEF1D) [21,20,22] as well as ribosomal proteins (such as s15, s20, SA) [21,20]. The kinase-activating LRRK2 mutant G2019S is reported to increase translation activity in *Drosophila* [23,21] using both direct and indirect measures for overall translation activity and protein translation reporter systems. In a related mechanism, LRRK2 interacts with argonaute proteins [24], which are essential components of miRNA machinery regulating protein expression levels by inhibition of translation or degradation of mRNA [25,26]. Therefore, LRRK2 may intervene in regulating protein expression via multiple mechanisms.

Interestingly, several recessive PD genes such as *PRKN*, *PINK1* and *DJ-1*, which cause early onset forms of PD, are also reported to influence protein translation. DJ-1 is a redox sensitive chaperone protein, and disease mutant forms of DJ-1 lead to translational defects in *E. coli* [27] through altered association with ribosomal proteins and decreased translation accuracy. PINK1 is enriched in the mitochondria, and acts together with Parkin to regulate mitophagy, an essential process of cellular homeostasis [28,29]. Individually, PINK1 is shown to be required for a translation switch from cap-dependent to cap-independent translation during hypoxia [30]. More recently, PINK1 and Parkin have been shown in *Drosophila* to direct localized translation of nuclear encoded respiratory chain complex mRNAs on the mitochondrial surface [31]. This is accomplished in mitochondria by the displacement of translation repressors, as well as the recruitment of translation activators such as eIF4G. Inversely, the inhibition of EIF4E in *Drosophila* influences parkin mutant phenotypes, suggesting a mutual interaction between parkin and protein translation [32].

#### *Clinical observations*

Additional clues on the importance of protein translation in PD come from observations in disease samples. Unbiased transcriptomics techniques to identify disease-specific cellular signatures and signaling pathways modified in disease recently showed a deregulation of pathways related to the translation initiation factor eIF2 in patient peripheral blood mononuclear cells (PBMCs) of both sporadic and genetic PD patients [33]. Of note, repression of translation by eIF2 signaling has been demonstrated in Prion disorder and postulated to occur in other prion-like disorders [34]. Indeed, eIF2 $\alpha$  phosphorylation is reported to be increased in PD and is closely associated with aggregates of  $\alpha$ -synuclein, indicating ER stress induced by misfolded proteins and increased protein translation through the unfolded protein response [35,36]. In addition, a recent study combining transcriptomics and proteomics of three PD related brain regions (substantia nigra, striatum and frontal cortex) to compare transcriptomes and proteomes in health and disease shows a subset of protein abundance changes largely independent of mRNA changes [37]. The study indicated that protein expression regulation mechanisms are active in PD. Control of translation in PD cases may also rely on preferential usage of an extended transcript of *SNCA*, the gene encoding  $\alpha$ -synuclein. Alpha-synuclein is a major component of Lewy bodies, a defining PD neuropathological feature, and  $\alpha$ -synuclein expression levels [38,39] and misfolding are considered key events in PD pathology [40]. The preferential usage of a long *SNCA* transcript is reported to favor  $\alpha$ -synuclein accumulation and subcellular localization [41], suggesting that translation of specific transcripts can contribute to pathological events. A role for  $\alpha$ -synuclein as an effector of toxicity dependent on protein translation is consistent with the notion that  $\alpha$ -synuclein is a key toxic protein factor downstream of other PD proteins in the pathological pathway [42].

#### *Potential mechanisms of protein translation deregulation in PD*

While these studies intuitively point to the hypothesis that deregulation of protein translation can lead to disease states, the question remains as to the precise mechanisms involved. In comparison to transcription, which is tightly regulated by transcriptional regulators that carry out a very specific program of gene expression by interacting with regulatory sequences in genes such as promoter regions, the mechanisms governing expression programs at the translational level are less evident. Nevertheless, recent studies show that protein translation factors can affect expression of specific sets of proteins that mediate specific phenotypes. For instance, several studies show a link between protein translation and cellular survival and proliferation. Persistent formation of the eIF4F complex (composed of several eIF4 proteins, including eIF4E, eIF4G and eIF4A) is observed in melanoma, colon and thyroid cancer cell lines, leading to resistance to anti-cancer therapies which can be reversed when the eIF4F complex is disrupted [43]. In other studies, eIF4G1 has been shown to selectively promote the translation of mRNAs involved in survival and DNA damage response in the

presence of DNA damage stress [44]. Specific links between changes in protein translation and phenotypes are also observed in the brain, for example when levels of eIF4E increase in mice, causing increased cap-dependent translation, synaptic and behavioural aberrations reminiscent of autism are observed [45]. This finding is in line with another study showing that eIF4E mediated synthesis of neuroligins is crucial to maintaining the balance of excitatory to inhibitory synaptic inputs in the central nervous system and its deregulation leads to autism spectrum disorders [46].

Next to protein translation complexes themselves, several other players are important in regulating protein translation such as signaling cascades regulating translation factors, enzymes influencing ribosome biogenesis, and miRNAs. Translation factor activity is regulated by cellular signaling processes (reviewed in [47]), such as signaling cascades mediated by PI3K, mTOR or MAPKs which modulate activity of general translation factors or factors influencing transport or stability of mRNAs. For example, the mTOR pathway, which responds to growth factors and is deregulated in PD and PD models [33,48], controls several translation proteins such as the eIF4E inhibitor eIF4E binding protein (4E-BP), the kinase of eEF2 or the ribosomal kinase S6K [47–49]. In this way, external stimuli such as growth factors or mitogens can affect protein translation activity. Ribosome methylation is a key step in ribosome biogenesis [50], which is favored by the interaction of the methyltransferase DNMT1, the ribosomal protein RPL26 and the ribosomal RNA 5.8S [51]. Interestingly, a decrease in DNMT1 has been observed in PD brains [52] and  $\alpha$ -synuclein has been reported to inhibit DNMT1 by sequestering it away from its substrates [53]. The series of additional factors regulating protein translation also includes miRNAs. Given the involvement of some PD associated proteins in microRNA biogenesis, part of the translation control mediated by these proteins may occur through the regulation of translation by microRNAs [25].

#### *Outstanding questions and issues on translational disruption in PD*

The protein translation hypothesis for PD pathogenesis suggests that survival of dopaminergic neurons is linked to subsets of proteins whose expression is altered at the translational level by functional networks involving proteins implicated in PD (Figure 2). However, at present the precise mechanisms linking PD associated proteins to changes in translation of specific proteins are incompletely understood. To begin to address this issue, mechanistic studies of the role of PD proteins in protein translation process are warranted. EIF4G1 is a direct actor in protein translation and future studies may then focus on its role in the protein translation complex and how this is affected by disease mutations. LRRK2 both interacts with and phosphorylates protein translation proteins, suggesting that LRRK2 may either play a direct role in protein translation complexes or an upstream regulatory role for protein translation proteins or both. For proteins of recessive PD genes



such as *PINK1* or *PRKN*, initial data suggests that there may be a specific role for these proteins in regulating translation on mitochondria. As precise mechanisms of PD-associated proteins in protein translation emerge, it will be interesting to examine differences between tissues and cell types.

Another key experiment is to use unbiased screening approaches to identify those subsets of proteins that are altered at the protein translation level by PD proteins, such as via correlating transcriptomics with proteomics data or screening of transcript usage in polysomes [54]. This information will help understand downstream cellular effects of PD protein mediated alterations in protein translation. At present it is not known if these effects are very specific, such as a change in expression of a small set of proteins, or more diffuse, such as an overall increase or overall decrease in protein translation levels. It should also be noted that besides altering expression of proteins, the action of PD proteins in protein translation might also affect protein folding, which begins during translation and is affected by translation rates [55]. Therefore special attention may be required to monitor protein folding in nascent proteins in the presence of disease forms of PD proteins, in particular for two main suspects of misfolding in PD,  $\alpha$ -synuclein and tau.

In addition to investigating the role of PD proteins in translation in experimental systems, it is necessary to verify findings of translation deregulation by PD proteins in the disease itself. In principle, findings from experimental systems can be tested in disease in a rather straightforward way. For instance, protein expression changes identified in protein translation profiling experiments in cellular and animal models can be tested in patient samples. These experiments have the potential of yielding interesting new biomarkers of early stages of disease, detectable either through biochemical assays or via non-invasive imaging. Confirmed biomarkers have the potential to be used as diagnostics for early detection of disease or as markers to monitor disease progression. Such markers are currently lacking in the field, yet are needed given that research on PD proteins has begun to move PD protein targeting therapies to late stage pre-clinical and early-stage clinical research [56–58].

Finally, the question of the druggability of protein translation poses itself. Many of the proteins mentioned above are enzymes, such as kinases or ubiquitin ligases, for which small compound modulators exist [59,60]. Whether these classes of compounds can rectify protein translation deregulation caused by PD proteins will be an important issue to address. Some small molecules targeting protein translation complexes themselves also exist. For example, small molecules targeting eIF4G1 have been developed as potential therapeutics in multiple myeloma; pharmacological inhibition of eIF4G1 causes a reduction in the expression of its molecular targets

[61] and increases cytotoxicity in breast cancer [62]. Other compounds targeting translation factors show that translation downstream of eIF2 can be pharmacologically restored in prion diseased mice [63]. Such recent advances show the feasibility of targeting protein translation processes with small molecules. Outstanding questions on protein translation deregulation in PD are summarized in Box 1.

#### *Concluding remarks*

At present, clinicians and patients lack diagnostic and therapeutic tools for early stage PD. Recent converging findings reveal a pressing need for in depth study of protein translation in PD. The hypothesis of translational deregulation as a key toxic mechanism in PD is a potential game changer in the field of the molecular pathogenesis of PD, with implications for the development of PD diagnostics and disease-modifying therapies.

**Box 1: Outstanding questions**

- Do PD proteins deregulate protein translation in disease?
  - Do EIF4G1 mutants deregulate protein translation?
  - Do LRRK2 mutants alter protein translation complexes?
  - How do PINK1 and PRKN mutants affect mitochondrial translation?
  - Which other PD proteins are involved in pathogenic deregulation of protein translation?
  - How do PD proteins intervene in other processes regulating protein translation such as the miRNA pathway?
- What is the relationship between PD proteins at the level of protein translation processes?
- What is the best model to study effects of PD proteins on protein translation? Do findings in invertebrate models translate to mammalian models? Do findings in animal models translate to the clinical situation?
- Do PD proteins regulate the expression of a specific subset of proteins through their deregulation of the protein translation processes?
  - If yes, can this subset of proteins be exploited as disease biomarkers?
- Does the role of PD proteins in protein translation affect dopamine neuron health?
- Does the role of PD proteins in protein translation influence the appearance of  $\alpha$ -synuclein toxic species and Lewy body formation?
- Is protein translation a therapeutic target for PD?
  - If so, what are possible targeting approaches?

## **Glossary**

**5' cap:** The chemical modification at the 5' end of mRNAs, often in the form of a modified guanine nucleotide, which stabilizes the mRNA. The 5' cap is required for efficient nuclear export as well as for the recruitment of translation initiation complexes in cap-dependent protein translation.

**Cap-dependent translation:** protein translation that involves recruitment of a translation initiation complex to the 5' cap structure of the mRNA to be translated.

**Cap-independent translation:** protein translation that involves recruitment of a translation initiation complex to a structure of the mRNA to be translated other than the 5' cap structure.

**Cap-Independent Translation Element (CITE):** a sequence in mRNA that mediates translation initiation without requiring the 5' cap.

**Eukaryotic elongation factor (eEF):** protein that is in complex with ribosomal proteins and the mRNA being translated which contributes to the synthesis of the polypeptide chain.

**Eukaryotic initiation factor (eIF):** protein which acts in complex with ribosomal proteins and other eIFs during the protein translation initiation phase, which includes recruitment of ribosomes to mRNA to be translated and scanning for the start codon.

**Genetic linkage study:** study of the co-segregation of a trait (such as disease) and an allele (such as mutation of a gene) in families. Linked genes are dominant when a single copy of the allele confers disease or recessive when disease occurs only when alleles are present on both chromosomes.

**Genome-wide association study (GWAS):** study searching across the entire genome for single nucleotide polymorphisms (SNPs) that vary between populations (such as of a population of PD patients compared to a matched control population). Polymorphisms enriched in disease populations point to candidate risk factors for that disease.

**Internal Ribosomal Entry Site (IRES):** a nucleotide sequence in mRNA that allows translation initiation to occur in the middle of the mRNA, as opposed to at the 5' cap.

**Mitochondrial translation:** protein translation occurring in mitochondria or on mitochondrial membranes. A very limited number of proteins are translated in mitochondria, mostly proteins that are components of the oxidative phosphorylation pathway. Mitochondrial translation is distinct from protein translation occurring in the cytoplasm and across the membrane of the endoplasmic reticulum, where the majority of cellular proteins are translated. Mitochondrial translation occurs for mitochondrially encoded genes within mitochondria but can also occur for nuclear encoded genes at the surface of mitochondrial outer membranes.

**Ribosome:** ribonucleoprotein complexes which serve as the protein synthesis machinery, linking amino acids together in polypeptide chains dictated by successive codon triplets of messenger RNA. Ribosomes act in conjunction with protein translation factors to fulfill the different steps of protein synthesis.



**Table 1: Overview of PD-associated genes reported to play a role in protein translation**

Gene	Protein encoded	Genetic linkage (locus, mLOD score)	Genome-Wide Association	Mendelian inheritance	Link to protein translation
<i>SNCA</i>	Alpha-synuclein	Yes (PARK1/4, mLOD 6.0)	Yes	Autosomal dominant PD	Alternate transcript usage in health and disease
<i>PRKN</i>	Parkin	Yes (PARK2, mLOD 7.7)		Early onset, autosomal recessive PD	Regulation of translation of respiratory chain complex mRNAs on mitochondrial outer membrane
<i>DJ-1</i>	DJ-1	Yes (PARK5, mLOD 4.2)		Early onset, autosomal recessive PD	Interacts with ribosomes in <i>E. coli</i>
<i>PINK1</i>	Pten-induced kinase 1	Yes (PARK6, mLOD 4.0)		Early onset, autosomal recessive PD	Regulation of translation of respiratory chain complex mRNAs on mitochondrial outer membrane
<i>LRRK2</i>	Leucine-rich repeat kinase 2	Yes (PARK8, mLOD 24.9)	Yes	Late onset, autosomal dominant PD	Interaction with protein translation proteins. G2019S disease mutant increases translation in <i>Drosophila</i>
<i>VPS35</i>	Vacuolar protein sorting-associated	Yes (PARK17, mLOD 1.4 <sup>a</sup> )		Late onset, autosomal dominant PD	Genetic interaction with EIF4G1

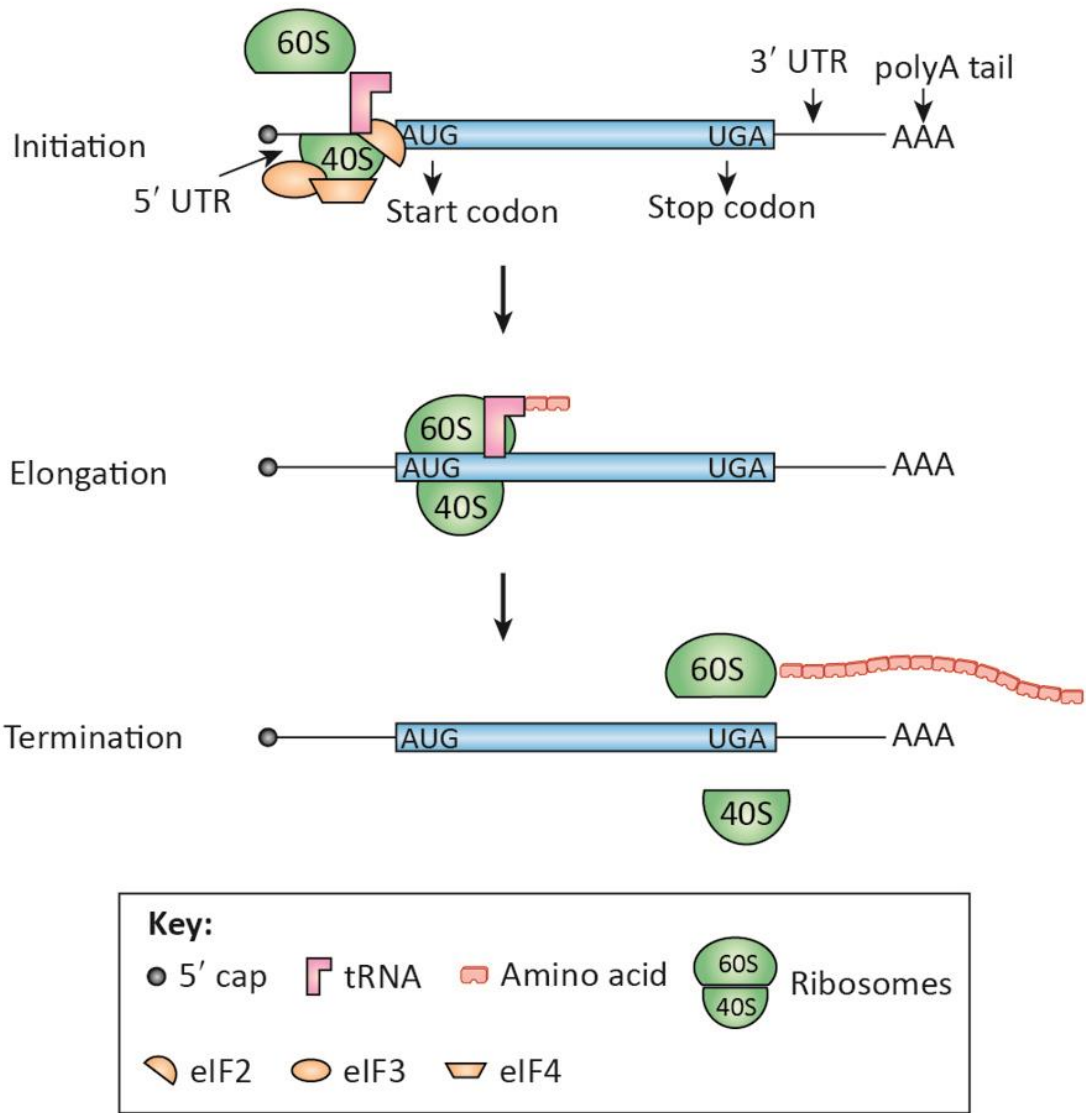
	protein 35				
<i>EIF4G1</i>	Eukaryotic translation initiation factor 4G1	Yes (PARK18, mLOD 3.0 <sup>b</sup> )		Late onset, autosomal dominant PD	EIF4G1 is a translation initiation factor

mLOD, maximum logarithm of odds, reviewed in reference [64].

<sup>a</sup>mLOD score does not reach the LOD threshold for conclusive linkage

<sup>b</sup>Rare variants with low penetrance

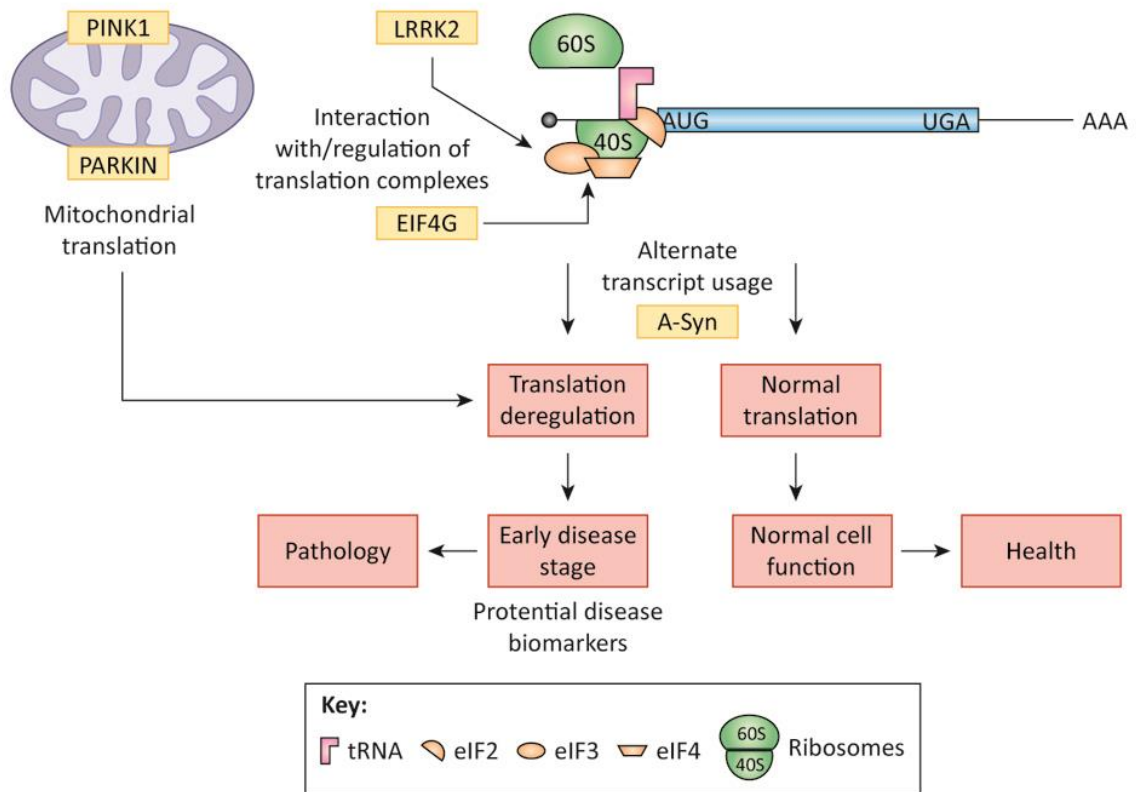
**Figure Legends**



**Figure 1: Schematic depiction of the three main steps of protein translation: initiation, elongation and termination.** Shown is an mRNA molecule including a coding sequence (in blue, including the AUG start and UGA stop codons), flanked by a 5'UTR with cap and polyadenylated 3'UTR. Ribosomal proteins, tRNAs and translation factors are recruited to the mRNA to mediate the different steps of the protein synthesis process, shown here for cap-dependent translation. At translation initiation eIF2 recruits tRNAs to the 40S ribosomal subunit. eIF3 and eIF4 initiation factors are also recruited to the 40S and this complex binds to the 5' cap structure of mRNAs and scans the 5' untranslated region (UTR) to find the start codon, at which point the 60S ribosomal subunit is recruited and initiation



factors are released. During the elongation phase, elongation factors (not depicted) are recruited to the ribosome and polypeptide chain synthesis occurs via the cycle of positioning the appropriate tRNA to the ribosome, forming the new peptide bond and shifting the ribosome complex by 1 codon on the mRNA chain. When the translation complex reaches the stop codon, elongation stops and the ribosome is released from the mRNA with the help of release factors (not depicted).



**Figure 2: Parkinson's disease proteins in protein translation.** Several PD proteins intervene in protein translation processes, schematically summarized here. LRRK2 is a regulator and interactor of protein translation factors; eIF4G1 is a translation factor itself. For alpha-synuclein (A-syn), the translation of alternate transcripts is correlated to normal cell function or pathology, while PINK1 and Parkin are reported to regulate protein translation of respiratory chain proteins at the mitochondrial outer membrane. While under non-disease circumstances, the activity of PD proteins in protein translation yields normal cellular function, it is hypothesized that this leads to pathology under disease conditions (deregulation such as the presence of a disease mutant form of PD protein or loss or gains of function).

## References

- 1 López-Lastra, M. *et al.* (2005) Protein synthesis in eukaryotes: The growing biological relevance of cap-independent translation initiation. *Biol. Res.* 38, 121–146
- 2 Shatsky, I.N. *et al.* Cap- and IRES-independent scanning mechanism of translation initiation as an alternative to the concept of cellular IRESs. , *Molecules and Cells*, 30. (2010) , 285–293
- 3 Smits, P. *et al.* (2010) Mitochondrial translation and beyond: processes implicated in combined oxidative phosphorylation deficiencies. *J. Biomed. Biotechnol.* 2010, 737385
- 4 Morimoto, K. *et al.* (2007) The functions of RPS19 and their relationship to Diamond-Blackfan anemia: a review. *Mol. Genet. Metab.* 90, 358–62
- 5 Kelleher, R.J. and Bear, M.F. (2008) The autistic neuron: troubled translation? *Cell* 135, 401–6
- 6 Scheper, G.C. *et al.* (2007) Translation matters: protein synthesis defects in inherited disease. *Nat. Rev. Genet.* 8, 711–23
- 7 Van der Knaap, M.S. *et al.* (2002) Mutations in each of the five subunits of translation initiation factor eIF2B can cause leukoencephalopathy with vanishing white matter. *Ann. Neurol.* 51, 264–70
- 8 Satterfield, T.F. and Pallanck, L.J. (2006) Ataxin-2 and its Drosophila homolog, ATX2, physically assemble with polyribosomes. *Hum. Mol. Genet.* 15, 2523–32
- 9 Magaña, J.J. *et al.* (2013) Spinocerebellar ataxia type 2: Clinical presentation, molecular mechanisms, and therapeutic perspectives. *Mol. Neurobiol.* 47, 90–104
- 10 Nusch, M. and Eckmann, C.R. (2013) *Germ Cell Development in C. elegans*, 757
- 11 Kozlov, G. *et al.* (2010) Structural basis of binding of P-body-associated proteins GW182 and ataxin-2 by the Mlle domain of poly(A)-binding protein. *J. Biol. Chem.* 285, 13599–13606
- 12 Kaehler, C. *et al.* Ataxin-2-Like Is a Regulator of Stress Granules and Processing Bodies. , *PLoS ONE*, 7. (2012)
- 13 Van de Loo, S. *et al.* (2009) Ataxin-2 associates with rough endoplasmic reticulum. *Exp. Neurol.* 215, 110–118

- 14 Nalls, M. a *et al.* (2014) Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. *Nat. Genet.* 056, 1–7
- 15 Chartier-Harlin, M.-C. *et al.* (2011) Translation initiator EIF4G1 mutations in familial Parkinson disease. *Am. J. Hum. Genet.* 89, 398–406
- 16 Fujioka, S. *et al.* (2013) Sequence variants in eukaryotic translation initiation factor 4-gamma (eIF4G1) are associated with Lewy body dementia. *Acta Neuropathol.* 125, 425–38
- 17 Sonenberg, N. and Dever, T.E. (2003) Eukaryotic translation initiation factors and regulators. *Curr. Opin. Struct. Biol.* 13, 56–63
- 18 Chartier-Harlin, M.-C. *et al.* (2013) , Functional Analyses of the eIF4G1 Mutations Linked to Parkinson's Disease. , in *Alzheimer's disease/Parkinson's disease (ADPD2013) conference, 2013, Firenze*, pp. 296
- 19 Dhungel, N. *et al.* (2015) Parkinson's Disease Genes VPS35 and EIF4G1 Interact Genetically and Converge on  $\alpha$ -Synuclein. *Neuron* 85, 76–87
- 20 Reyniers, L. *et al.* (2014) Differential protein-protein interactions of LRRK1 and LRRK2 indicate roles in distinct cellular signaling pathways. *J. Neurochem.* DOI: 10.1111/jnc.12798
- 21 Martin, I. *et al.* (2014) Ribosomal protein s15 phosphorylation mediates LRRK2 neurodegeneration in Parkinson's disease. *Cell* 157, 472–85
- 22 Dorval, V. and Hébert, S.S. (2012) LRRK2 in Transcription and Translation Regulation: Relevance for Parkinson's Disease. *Front. Neurol.* 3, 12
- 23 Imai, Y. *et al.* (2008) Phosphorylation of 4E-BP by LRRK2 affects the maintenance of dopaminergic neurons in *Drosophila*. *EMBO J.* 27, 2432–43
- 24 Gehrke, S. *et al.* (2010) Pathogenic LRRK2 negatively regulates microRNA-mediated translational repression. *Nature* 466, 637–41
- 25 Bartel, D.P. (2009) MicroRNAs: Target Recognition and Regulatory Functions. *Cell* 136, 215–233
- 26 Cipolla, G. a (2014) A non-canonical landscape of the microRNA system. *Front. Genet.* 5, 1–6

- 27 Kthiri, F. *et al.* (2010) Translational defects in a mutant deficient in YajL, the bacterial homolog of the parkinsonism-associated protein DJ-1. *J. Bacteriol.* 192, 6302–6
- 28 Narendra, D.P. *et al.* (2010) PINK1 is selectively stabilized on impaired mitochondria to activate Parkin. *PLoS Biol.* 8, e1000298
- 29 Geisler, S. *et al.* (2010) PINK1/Parkin-mediated mitophagy is dependent on VDAC1 and p62/SQSTM1. *Nat. Cell Biol.* 12, 119–131
- 30 Lin, W. *et al.* (2014) Loss of PINK1 attenuates HIF-1 $\alpha$  induction by preventing 4E-BP1-dependent switch in protein translation under hypoxia. *J. Neurosci.* 34, 3079–89
- 31 Gehrke, S. *et al.* (2015) Article PINK1 and Parkin Control Localized Translation of Respiratory Chain Component mRNAs on Mitochondria Outer Membrane. *Cell Metab.* 21, 95–108
- 32 Ottone, C. *et al.* (2011) Diminution of eIF4E activity suppresses parkin mutant phenotypes. *Gene* 470, 12–9
- 33 Mutez, E. *et al.* (2014) Involvement of the immune system, endocytosis and EIF2 signaling in both genetically determined and sporadic forms of Parkinson's disease. *Neurobiol. Dis.* 63, 165–70
- 34 Moreno, J. a *et al.* (2012) Sustained translational repression by eIF2 $\alpha$ -P mediates prion neurodegeneration. *Nature* 485, 507–11
- 35 Hoozemans, J.J.M. *et al.* (2007) Activation of the unfolded protein response in Parkinson's disease. *Biochem. Biophys. Res. Commun.* 354, 707–711
- 36 Hoozemans, J.J.M. *et al.* (2012) Activation of the unfolded protein response is an early event in Alzheimer's and Parkinson's disease. *Neurodegener. Dis.* 10, 212–5
- 37 Riley, B.E. *et al.* (2014) Systems-based analyses of brain regions functionally impacted in Parkinson's disease reveals underlying causal mechanisms. *PLoS One* 9, e102909
- 38 Chartier-Harlin, M.-C. *et al.* (2004) Alpha-synuclein locus duplication as a cause of familial Parkinson's disease. *Lancet* 364, 1167–9

- 39 Singleton, A.B. *et al.* (2003) alpha-Synuclein locus triplication causes Parkinson's disease. *Science* 302, 841
- 40 Visanji, N.P. *et al.* (2013) The prion hypothesis in Parkinson's disease: Braak to the future. *Acta Neuropathol. Commun.* 1, 2
- 41 Rhinn, H. *et al.* (2012) Alternative  $\alpha$ -synuclein transcript usage as a convergent mechanism in Parkinson's disease pathology. *Nat. Commun.* 3, 1084
- 42 Taymans, J.-M. and Cookson, M.R. (2010) Mechanisms in dominant parkinsonism: The toxic triangle of LRRK2, alpha-synuclein, and tau. *Bioessays* 32, 227–35
- 43 Boussemart, L. *et al.* (2014) eIF4F is a nexus of resistance to anti-BRAF and anti-MEK cancer therapies. *Nature* 513, 105–9
- 44 Badura, M. *et al.* (2012) DNA damage and eIF4G1 in breast cancer cells reprogram translation for survival and DNA repair mRNAs. *Proc. Natl. Acad. Sci. U. S. A.* 109, 18767–72
- 45 Santini, E. *et al.* (2013) Exaggerated translation causes synaptic and behavioural aberrations associated with autism. *Nature* 493, 411–5
- 46 Gkogkas, C.G. *et al.* (2013) Autism-related deficits via dysregulated eIF4E-dependent translational control. *Nature* 493, 371–7
- 47 Proud, C.G. (2007) Signalling to translation: how signal transduction pathways control the protein synthetic machinery. *Biochem. J.* 403, 217–34
- 48 Tain, L.S. *et al.* (2009) Rapamycin activation of 4E-BP prevents parkinsonian dopaminergic neuron loss. *Nat. Neurosci.* 12, 1129–35
- 49 Liu, S. and Lu, B. (2010) Reduction of protein translation and activation of autophagy protect against PINK1 pathogenesis in drosophila melanogaster. *PLoS Genet.* 6, 1–12
- 50 Kalpana Ghoshal, Sarmila Majumder, Jharna Datta, Tasneem Motiwala, Shoumei Bai, Sudarshana M. Sharma, Wendy Frankel, and S.T.J. (2012) Role of Human Ribosomal RNA (rRNA) Promoter Methylation and methyl-CpG-binding protein MBD2 in the suppression of rRNA gene expression. *Changes* 29, 997–1003

- 51 Mastroeni, D. *et al.* (2011) Epigenetic changes in Alzheimer's disease: decrements in DNA methylation. *31*, 2025–2037
- 52 Jowaed, A. *et al.* (2010) Methylation regulates alpha-synuclein expression and is decreased in Parkinson's disease patients' brains. *J. Neurosci.* *30*, 6355–9
- 53 Desplats, P. *et al.* (2011) Alpha-synuclein sequesters Dnmt1 from the nucleus: a novel mechanism for epigenetic alterations in Lewy body diseases. *J. Biol. Chem.* *286*, 9031–7
- 54 Ingolia, N.T. (2014) Ribosome profiling: new views of translation, from single codons to genome scale. *Nat. Rev. Genet.* *15*, 205–13
- 55 O'Brien, E.P. *et al.* (2014) Understanding the influence of codon translation rates on cotranslational protein folding. *Acc. Chem. Res.* *47*, 1536–44
- 56 AlDakheel, A. *et al.* (2014) Pathogenesis-Targeted, Disease-Modifying Therapies in Parkinson Disease. *Neurotherapeutics* *11*, 6–23
- 57 Olanow, C.W. and Schapira, A.H. V (2013) Therapeutic prospects for parkinson disease. *Ann. Neurol.* *74*, 337–347
- 58 Taymans, J. (2014) Can the increasing number of newly developed leucine-rich repeat kinase 2 inhibitors validate or invalidate a potential disease-modifying therapeutic approach for Parkinson's disease? *Expert Opin. Ther. Pat.* *24*, 727–30
- 59 Nalepa, G. *et al.* (2006) Drug discovery in the ubiquitin-proteasome system. *Nat. Rev. Drug Discov.* *5*, 596–613
- 60 Vancaenenbroeck, R. *et al.* (2011) Kinases as targets for Parkinson's disease: from genetics to therapy. *CNS Neurol. Disord. Drug Targets* *10*, 724–40
- 61 Attar-Schneider, O. *et al.* (2014) Targeting eIF4GI translation initiation factor affords an attractive therapeutic strategy in multiple myeloma. *Cell. Signal.* *26*, 1878–1887
- 62 Yi, T. *et al.* (2014) 4EGI-1 targets breast cancer stem cells by selective inhibition of translation that persists in CSC maintenance , proliferation and metastasis. *Oncotarget* *5*, 6028–6037

- 63 Halliday, M. *et al.* (2015) Partial restoration of protein synthesis rates by the small molecule ISRIB prevents neurodegeneration without pancreatic toxicity. *Cell Death Dis.* 6, e1672
- 64 Verstraeten, A. *et al.* (2015) Progress in unraveling the genetic etiology of Parkinson disease in a genomic era. *Trends Genet.* 31, 140–149

# DISCUSSION



## 4. Discussion générale

Depuis sa découverte en 1817, les connaissances sur l'origine anatomopathologique et les mécanismes biochimiques et moléculaires de la MP ont grandement évolué. Cependant, alors que certains de ces mécanismes ont été identifiés, d'autres demeurent encore un mystère et expliquent en partie que la MP reste aujourd'hui une maladie incurable. La compréhension de ces mécanismes étant le point de départ de la recherche et l'identification de biomarqueurs et de cibles thérapeutiques, l'amélioration de leur connaissance est une problématique actuelle et d'une importance capitale pour une meilleure prise en charge de la MP. Afin d'apporter un éclairage sur les zones encore obscures de la physiopathologie moléculaire de la MP, nous avons eu recours à différentes méthodes d'analyse du transcriptome de cellules mononucléées périphériques sanguines dérivées de patients et de témoins appariés. Par l'étude des profils d'expression génique de ces cellules, ces travaux de thèse ont mis en exergue des variations quantitatives de gènes et transcrits impliqués dans des processus physiopathologiques connus validant ainsi la pertinence de la stratégie établie ainsi que d'autres variations plus novatrices. Celles-ci ont souligné l'existence de perturbations du métabolisme des ARNs qui pourraient constituer un nouveau mécanisme de la physiopathologie de la maladie et être source de biomarqueurs.

### 4.1. Études des mécanismes moléculaires de la maladie de Parkinson

L'utilisation de puces d'expression avait été requise lors d'une première analyse du transcriptome de cellules mononucléées périphériques sanguines de patients porteurs de la mutation G2019S de *LRRK2* avant mon arrivée dans le laboratoire du Dr. Marie-Christine Chartier-Harlin. Cette étude avait mis en exergue, chez les parkinsoniens, des dérèglements de la réponse inflammatoire, la réponse immunitaire, la prolifération et la mort cellulaire. Elle avait également révélé des variations d'expression de gènes impliqués dans la maturation des ARNs (incluant l'épissage) et leur traduction. La comparaison de ces données avec des données transcriptomiques de la littérature provenant de la substance noire de parkinsoniens, avait permis 1) de rendre compte de la pertinence de ces dérèglements dans la MP puisque des variations au niveau des gènes liés à la réponse inflammatoire, à la réponse immunitaire, à la prolifération et à la mort cellulaire ont également été observées au niveau cérébral<sup>279,297</sup>, 2) de valider les cellules mononucléées périphériques sanguines comme modèle d'étude capable de mimer les mécanismes physiopathologiques de la MP.

Sur la base de cette validation, nous avons choisi ces cellules pour effectuer différentes analyses du transcriptome afin de mettre en évidence des variations quantitatives à l'aide de puces

d'expression ainsi que des variations quantitatives couplées à des études de variations qualitatives par le biais du RNAseq. En associant ces deux méthodes, nous avons à nouveau mis en évidence des perturbations quantitatives de gènes impliqués dans les voies citées ci-dessus ainsi que dans l'endocytose, de la mitochondrie et de l'ubiquitination augmentant la pertinence de ces dérèglements dans la MP. De plus, nous avons confirmé les variations de gènes liés au métabolisme des ARNs et révélé des variations quantitatives et qualitatives de l'épissage. Le fait de retrouver ces mêmes dérèglements à l'aide d'outils différents et d'échantillons différents tend à confirmer l'existence de ces perturbations et révèle que ces défauts de l'épissage sont liés à la maladie. Ceci est d'autant plus vrai que nous avons montré que les cellules mononuclées périphériques sanguines étaient capables de refléter des perturbations moléculaires associées à la maladie. Ainsi, bien que ces résultats soient préliminaires, ils nous confortent dans notre choix d'étudier le transcriptome de cellules sanguines et nous invitent à poursuivre nos analyses par le biais de ces outils pour mieux comprendre les mécanismes de la maladie.

La pertinence de ces modifications que nous avons identifiées et leur implication dans la physiopathologie de la MP est d'autant plus grande qu'elles sont confirmées par les données transcriptomiques de cellules sanguines de parkinsoniens obtenues par d'autres équipes. En effet, Scherzer *et al.* avaient proposé en 2007, la première analyse du transcriptome de cellules sanguines totales dérivées de 50 patients *versus* 55 témoins dans le but d'identifier des marqueurs précoces. Cette étude avait montré des variations d'expression de 22 gènes entre les malades et les témoins, ces derniers étant impliqués dans l'ubiquitination, l'endocytose des transporteurs à la dopamine, les troubles mitochondriaux, l'ubiquitination et le métabolisme des protéines<sup>240</sup>. De même, par le biais de puces à exons, Soreq et ses collaborateurs avaient montré que des modifications de l'épissage s'opéraient dans les cellules mononuclées périphériques sanguines de patients ayant subi ou non une stimulation cérébrale profonde et que ces changements accompagnaient les bénéfices physiques conférés par le traitement<sup>298</sup> suggérant un lien entre ces changements moléculaires et la MP. Par l'analyse des profils d'expression de 4 parkinsoniens non traités à des stades très précoces (HY=1) et de 4 témoins appariés, Alieva *et al.* avaient également identifié, dans les cellules sanguines, des variations de processus cellulaires relatifs au système immunitaire, au transport cellulaire ainsi que de plusieurs gènes impliqués dans l'épissage et dans l'endocytose<sup>284</sup>. Le séquençage des ARNs de monocytes a montré des variations de la régulation de la production de cytokines (inflammation), et de la réponse au stress<sup>237</sup>. Des études du transcriptome par RNAseq de cellules sanguines de sujets porteurs de la mutation G2019S de *LRRK2* ont montré une perturbation de gènes liés aux interactions entre les neurones et la glie, des protéines de la guidance axonale et du métabolisme du glucose<sup>244</sup>. Enfin, des perturbations des voies de la réponse et des maladies inflammatoires ont été

retrouvées par l'analyse du profil d'expression des transcrits, à l'aide de puce, à partir de sang total de 34 patients porteurs de la mutation G2019S de *LRRK2* versus 32 sujets asymptomatiques dont la moitié sont porteurs de la mutation G2019S de *LRRK2*<sup>299</sup>. Ceci soutient l'implication de la réponse immunitaire dans les premières étapes des processus pathologiques.

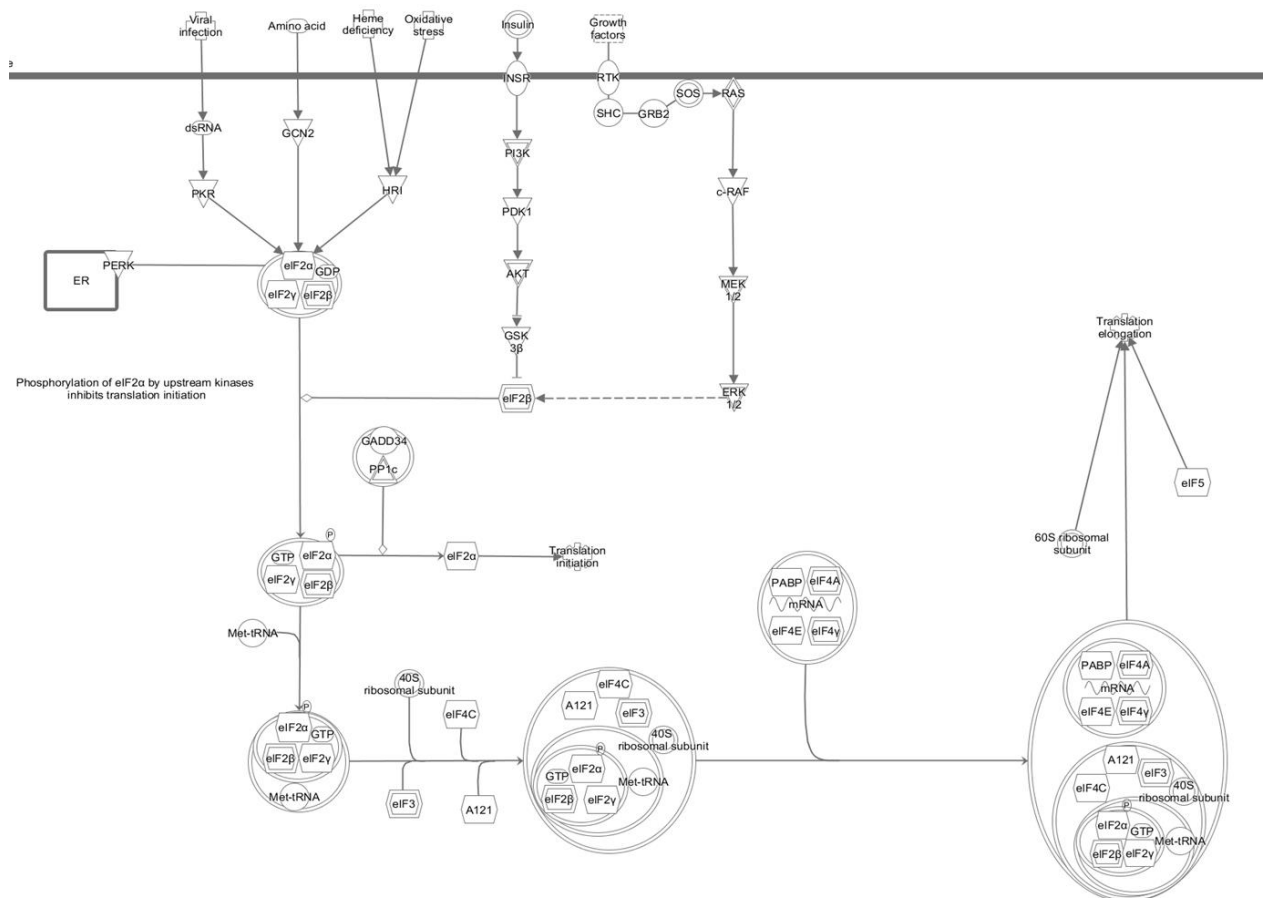
Outre la confirmation des données présentes à de multiples reprises dans la littérature, nos différentes analyses du transcriptome viennent compléter ces informations notamment concernant les défauts d'épissage qui eux n'ont pas autant été explorés. En effet, même si quelques rares équipes ont suggéré l'existence de tels dérèglements avant que nous commençons notre étude, l'ensemble de ces expertises a été basé sur l'utilisation de puces qui n'ont pas permis d'entrevoir toute l'étendue des modifications liés aux ARNs dans la MP.

#### 4.2. Importance de la voie de signalisation EIF2 dans la maladie de Parkinson : son lien avec la traduction et l'épissage

Notre analyse du transcriptome de cellules mononucléées périphériques sanguines de sujets porteurs de la mutation G2019S de *LRRK2* et de sujets sporadiques nous a permis d'identifier des voies moléculaires perturbées chez ces deux groupes de malades. Cette étude a contribué à l'identification de mécanismes qui pourraient expliquer les similitudes phénotypiques entre ces deux formes de la maladie ; d'une part une mutation très fréquente et d'autre part des cas sporadiques qui représentent la forme majoritaire de la maladie. Au-delà de l'identification de mécanismes partagés entre ces deux groupes, cette analyse avait pour but d'augmenter les chances d'identifier des perturbations moléculaires communes à un grand nombre de malades. Nous avons considéré cet aspect dans la mesure où le but final de l'amélioration des connaissances des mécanismes de la MP est l'identification de biomarqueurs diagnostiques et de cibles thérapeutiques permettant de détecter et de traiter le plus grand nombre de cas.

Ainsi, cette étude a mis en évidence des changements d'expression de gènes principalement impliqués dans la réponse immunitaire, les mécanismes d'endocytose et la voie de signalisation EIF2. Nous avons accordé une attention toute particulière à cette voie (figure 32) qui s'est avérée être la voie canonique, commune à ces deux formes, la plus significativement perturbée. La mise en évidence de perturbations de cette voie essentiellement connue pour son rôle dans la traduction, suggèrent que des modifications de la synthèse protéique pourraient avoir lieu dans la MP. Bien que des perturbations de cette voie n'aient pas été retrouvées au cours de notre étude présentée dans le chapitre 2, l'implication de la traduction est appuyée par l'identification de perturbations de la voie « Translation factor activity, RNA binding » chez les cas sporadiques et chez les parkinsoniens

porteurs de mutation de l'ATXN2 ainsi que la possible implication des protéines de la famille PABP dans l'établissement d'un phénotype parkinsonien. Nos analyses de RNAseq ont également suggéré des perturbations de la traduction par l'identification de changements de jonctions entre exons au sein de gènes impliqués dans les voies de signalisation « régulation de eIF4 et p70S6K » et à nouveau la « voie de signalisation EIF2 ». Cette dernière peu étudiée jusqu'alors dans le cadre de la MP est dérégulée dans plusieurs de nos analyses et pourrait constituer un nouveau mécanisme physiopathologique de la maladie puisqu'à l'heure actuelle seuls de rares exemples font état de perturbations de la synthèse protéique dans la MP.



**Figure 32 : Voie de signalisation EIF2 provenant d'Ingenuity Pathways Analysis**

La traduction est un processus qui s'effectue en 3 étapes distinctes : l'initiation, l'élongation et la terminaison. Chacune de ces étapes est finement contrôlée bien que l'initiation constitue l'étape limitante la plus régulée. La synthèse protéique commence par l'interaction de la protéine eIF4F (eukaryotic initiation factor 4F) avec la coiffe en 5' de l'ARNm. eIF4F est un complexe formé de 3 sous-unités: eIF4E (eukaryotic initiation factor 4E) qui est une protéine de liaison à la coiffe, eIF4A qui est une RNA hélicase et la protéine eIF4G qui permet le recrutement des ribosome à travers son interaction avec eIF3 qui interagit avec les ribosomes. La sous unité 40S du ribosome va se déplacer le long de la région 5'UTR de l'ARNm, à la recherche du codon ATG d'initiation de la traduction et permettre l'assemblage de la sous-unité 80S du ribosome et initier la réaction d'élongation et la synthèse des protéines

Pour tenter de comprendre le potentiel de cette voie comme mécanisme physiopathologique de la maladie, il est important de s'intéresser à son rôle dans la cellule.

La protéine EIF2 fait partie du complexe d'initiation de la traduction. La phosphorylation de sa sous-unité alpha (EIF2 $\alpha$ ) est un mode de régulation de la traduction très étudié et souvent appliqué en réponse à un stress cellulaire. Le choc thermique, le choc osmotique, l'irradiation par les UV, l'hypoxie ou encore le stress du réticulum endoplasmique (RE) ont été montrés comme inducteur de la phosphorylation de EIF2 $\alpha$  ; ces deux derniers étant rapportés dans la MP. La phosphorylation de EIF2 $\alpha$  entraîne une inhibition de la synthèse protéique globale au profit de protéines permettant à la cellule de répondre à ces stress<sup>300</sup>. Elle fait intervenir de nombreuses kinases telles que HRI, PKR, PERK, GCN2 et mTOR qui phosphorylent EIF2 $\alpha$  sur la sérine 51<sup>301,302</sup>. mTOR est un complexe protéique qui se situe en aval de la cascade de signalisation PI3K/Akt. Cette voie joue un rôle dans la survie et la croissance cellulaire ainsi que dans l'apoptose. Elle est activée dans de nombreux cancers et promeut la prolifération des cellules. A l'inverse, l'inhibition de cette voie a été montrée comme étant inductrice de la mort cellulaire de façon dépendante de la phosphorylation de EIF2 $\alpha$ <sup>303,304</sup>. De façon intéressante, des perturbations de la voie PI3K/Akt ont été retrouvées dans l'ensemble de nos données du transcriptome issues de puces d'expression et également dans nos données de RNAseq. Un lien entre les perturbations de ces deux voies est possible, ce d'autant que des perturbations de la voie PI3K/Akt sont retrouvées dans les neurones dopaminergiques de la *SNpc* de parkinsoniens<sup>227,279</sup> et que l'analyse des cerveaux *post-mortem* de parkinsoniens indique une augmentation de EIF2 $\alpha$  phosphorylée par rapport aux témoins<sup>305</sup>. L'augmentation de la phosphorylation d'EIF2 $\alpha$  dans les neurones dopaminergiques de parkinsoniens est associée à une augmentation de la phosphorylation de la kinase PERK. L'activation de PERK par phosphorylation témoigne de l'induction de la réponse UPR (Unfolded Protein Response). La voie UPR est activée en réponse à un stress du RE, site de repliement et des modifications post-traductionnelles des protéines destinées à la membrane plasmique ou aux organelles. Il est sensible à des perturbations de l'homéostasie cellulaire qui sont conduites par des stress tels que la déplétion calcique, la privation en glucose, l'hypoxie ou le mauvais repliement des protéines. L'activation de l'UPR induit une activation de PERK qui va phosphoryler eIF2 $\alpha$  et inhiber la traduction générale pour permettre à la cellule de prendre en charge la surcharge protéique. L' $\alpha$ -synucléine, dont l'agrégation est observée chez la majorité des parkinsoniens, est une des protéines qui s'accumule dans le RE et induit la réponse UPR dans les lignées cellulaires humaines embryonnaires de foie HEK-239T, des lignées cellulaires dopaminergiques de neuroblastome SY-5Y ainsi que dans le cerveau de souris<sup>306</sup>. Par ailleurs, l'augmentation de la phosphorylation de PERK, retrouvée dans les neurones dopaminergiques de parkinsoniens, est associée à une augmentation de l' $\alpha$ -synucléine suggérant un

lien entre l'accumulation de l' $\alpha$ -synucléine et le stress du RE<sup>305,307</sup> pouvant être responsable de l'induction de la réponse UPR et de perturbations de la voie EIF2.

Cette voie est également régulée par la phosphorylation de la protéine 4E-BP. Cette protéine interagit avec EIF4E bloquant ainsi son interaction avec le complexe initiateur de la traduction et inhibant cette dernière. La phosphorylation de 4E-BP permet de libérer EIF4E et d'initier la traduction. Chez l'homme, 4E-BP contient 6 sites phosphorylables au niveau des acides aminés thréonines 37, 46, 70 et sérines 65, 83 et 112. LRRK2 est capable de phosphoryler 4E-BP au niveau des thréonines 37 et 46 *in vitro* et *in vivo*<sup>139</sup>. Chez la drosophile, la surexpression de l'orthologue de LRRK2 sauvage ou porteur de la mutation G2019S conduit à une hyperphosphorylation de 4E-BP<sup>139</sup>. Le développement d'un modèle de drosophile invalidé pour le gène *4E-BP* a montré que la perte de la protéine entraînait une diminution de 25% de la durée de vie des drosophiles par rapport aux drosophiles contrôles<sup>308</sup>. Il se pourrait que l'action de LRRK2 sur l'activation de la voie EIF2 joue un rôle également dans la survie rejoignant l'idée du rôle des perturbations de cette voie dans la mort des neurones dopaminergiques. L'exposition de drosophiles transgéniques exprimant la forme mutée G2019S de *LRRK2* à un stress oxydant induit par le paraquat ou l' $H_2O_2$  augmente leur sensibilité à ce stress et entraîne une réduction du nombre de neurones dopaminergiques et de la quantité de dopamine dans leur cerveau<sup>139</sup>. Le stress oxydant induit par la perte de fonction de PINK1 dans des fibroblastes embryonnaires de souris est associée à une augmentation de la phosphorylation de 4E-BP1, une augmentation de la traduction et à la mort des cellules portant une déplétion de PINK1<sup>309</sup>. L'ensemble de ces données appuie à nouveau l'implication de la voie de signalisation EIF2 dans la dégénérescence des neurones dopaminergiques chez les parkinsoniens, ce d'autant que des variations d'expression de gènes liés à cette voie sont retrouvées au niveau des neurones dopaminergiques des patients atteints de la MP<sup>279,297</sup>. Van de Berg *et al.* ont observé une activation de cette voie dans la *SNpc* de sujets dit « incidents » c'est-à-dire des sujets présentant des agrégats cérébraux d' $\alpha$ -synucléine dans les régions associées à la maladie et ce sans manifester de signes cliniques. Ces sujets sont assimilés à des stades précoces de la maladie<sup>310</sup> corroborant l'hypothèse de l'implication de la voie de signalisation EIF2 dans la genèse de la MP.

En analysant plus particulièrement les gènes contenus dans la voie de signalisation EIF2, différemment exprimés à la fois chez les patients porteurs de la mutation G2019S de *LRRK2* et chez les sujets sporadiques, nous avons observé que ces gènes étaient des gènes de protéines ribosomiques (RPL19, RPL21, RPL27A, RPL4, RPL9, RPS2, RPS9). Des études faites chez *C. elegans* et dans des lignées cellulaires épithéliales humaines HeLa et Calu 6 ont montré que les protéines ribosomiques étaient capables de réguler l'épissage de leur propre ARN. Le mécanisme n'est pas

clairement identifié mais il semblerait que ces protéines puissent interagir avec des facteurs d'épissage de type hnRNPs pour les recruter au niveau des sites d'épissage<sup>311,312,313</sup>. Le dérèglement de gènes appartenant à cette famille de protéines laisse penser que la perturbation de la voie de signalisation EIF2 témoignerait aussi en partie d'une modulation de l'épissage.

Certains de ces défauts d'épissage pourraient être la conséquence d'un manque de disponibilité des facteurs d'épissage pour effectuer leur action, ce d'autant que la protéine ataxine-2 a été montrée comme étant capable de séquestrer TDP43 via ses expansions de triplets CAG assurant ainsi un maintien de TDP43 dans le cytoplasme pouvant être à l'origine de certains troubles de l'épissage<sup>145</sup>. De façon intéressante, *PABPC4* que nous avons vu comme étant dérégulé chez les parkinsoniens est également dérégulé dans les neurones moteurs de patients atteints de SLA où sa protéine est séquestrée au niveau des foci<sup>314</sup>. Par l'analyse du transcriptome de patients présentant des expansions nucléotidiques dans le gène *C9orf72* (responsables de certaines cas de SLA), Cooper-Knock *et al.* ont révélé que 89% des gènes codant des protéines liées à l'épissage étaient séquestrées au niveau des foci et dérégulés dans les neurones moteurs dérivés de patients (contre 20% au niveau des lignées lymphoblastoïdes) induisant des défauts d'épissage<sup>314</sup>. Sachant que la SLA et la MP présentent des similitudes, notamment concernant les expansions de triplets CAG dans *ATXN2*, il est possible que la séquestration des protéines de l'épissage soit un mécanisme commun à ces deux maladies.

La régulation de l'épissage fait intervenir plusieurs signaux extracellulaires qui sont transmis au noyau par l'intermédiaire de cascade de signalisation. La voie PI3K/Akt est une des voies qui permet de réguler l'épissage notamment par la phosphorylation des effecteurs (paragraphe 1.1.5.4). Plusieurs exemples de la littérature montrent son implication dans la régulation de l'épissage de *CD44* dans les cellules T<sup>315,316</sup> mais aussi de la fibronectine<sup>317</sup> ou de *XBPI*<sup>318</sup> notamment par l'activation du facteur d'épissage SF2 (protéine SR) par phosphorylation via la protéine mTOR<sup>319</sup>. Sachant que la phosphorylation est un moyen de modifier la localisation subcellulaire des facteurs d'épissage (paragraphe 1.1.5.4.), les perturbations de la voie PI3K/Akt que nous avons observées dans l'ensemble de nos études pourraient être à l'origine de modifications qualitatives de l'épissage par modification de la localisation de certains facteurs dans la cellule. En traitant des lignées cellulaires dopaminergiques SY-5Y avec du paraquat, Vivarelli *et al.* ont révélé que l'induction d'un stress oxydant conduisait à une phosphorylation de la protéine SRPK2, une protéine kinase qui régule l'épissage par phosphorylation des protéines SR. La phosphorylation de SRPK2 induit un changement de sa localisation du cytoplasme vers le noyau qui est alors capable d'hyperphosphoryler ses protéines SR cibles. A l'aide de minigène, ils ont montré que cette hyperphosphorylation induisait

des variations de l'épissage et suggéré que les défauts de ce processus dans la MP pourraient être associés à des modifications de la disponibilité de ses régulateurs<sup>320</sup>.

#### 4.3. Perturbations du métabolisme des ARNs : Implication des mécanismes de dégradation

L'étude du transcriptome intègre aussi les ARNs qui doivent être éliminés par la cellule. L'épissage alternatif est connu depuis longtemps pour être étroitement lié au nonsense-mediated decay (NMD), un mécanisme qui contrôle la durée de vie des ARNm et qui sert également de système de contrôle qualité de la cellule. Il permet ainsi l'élimination des ARNs aberrants. Ce lien réside dans la capacité de l'épissage à générer des codons stop prématures (ou codon non sens) qui vont induire un arrêt prématuré de la traduction. Les événements d'épissage alternatif de type inclusion/exclusion d'exons ou les épissages alternatifs de type 5' ou 3' peuvent entraîner des changements de cadre de lecture faisant apparaître des codons stop précoces.

Les effecteurs du NMD sont les protéines UPF1, UPF2 et UPF3 et se fixent à l'ARN à dégrader en interagissant à la fois avec les facteurs de terminaison de la traduction comme eRF1 et eRF3 ainsi qu'avec des enzymes de dégradation des ARNs. Environ 10% des ARNm totaux sont régulés par le NMD bien que certains aspects de sa régulation soient encore méconnus. Toutefois, plusieurs propositions ont été faites concernant le mécanisme par lequel le NMD détermine ses cibles :

- selon la position du codon stop prémature par rapport aux jonctions exons-exons, mode de reconnaissance le plus connu chez l'homme. Les codons prématures sont reconnus notamment en fonction de la distance qui les sépare de la liaison exons-exons : un codon stop situé à environ 50-55nt d'une liaison entre exons est une cible potentielle. Les jonctions exons-exons sont marquées par la présence des complexes EJC (exon junction complex) qui se situent à 20-24 nt en amont d'une jonction d'épissage. Ainsi, UPF3 va s'associer au complexe EJC dans le noyau et permettre le recrutement d'UPF2 qui va favoriser l'export de l'ARN dans le cytoplasme. La traduction va commencer puis lorsque le ribosome va atteindre un stop en amont d'un EJC, le recrutement d'UPF1 par relargage des facteurs de traduction va avoir lieu et permettre le recrutement d'UPF2 et UPF3 conduisant à la dégradation de l'ARN
- selon la présence d'introns dans les régions 3' non traduites
- selon la longueur de la région 3' non traduite de l'ARNm qui sépare, dans l'espace, le codon stop prémature de la queue polyA.
- ces codons prématures surviennent suite à des mutations, à des erreurs au cours de la transcription ou à des événements d'épissage alternatif.



Le NMD dépendant de la queue polyA implique les protéines de la famille PABP (polyA binding protein)<sup>321</sup>. Il existe 2 catégories de protéines PABPs chez les mammifères : les protéines nucléaires qui jouent un rôle dans la synthèse de la queue polyA et les protéines cytoplasmiques PABPC1, PABPC2, PABPC3, PABPC4 et PABPC5<sup>322</sup>. La capacité de PABPC1 à interagir avec la queue polyA et avec le complexe d'initiation de la traduction au niveau de la coiffe permet de rapprocher PABPC1 du codon stop prémature par la circularisation de l'ARN et donc de recruter le facteur Erf3 de terminaison de la traduction<sup>323,324</sup>. Le NMD s'enclenche lorsque le ribosome arrive au niveau d'un codon stop. Normalement ce codon constitue un signal de détachement du ribosome de l'ARNm en cours de traduction. Ce détachement se fait à l'aide du facteur de relargage eRF3 (eukaryotic release factor). Si le ribosome ne se décroche pas, le facteur Erf3 va recruter UPF1 et mettre en marche le mécanisme de dégradation. Bien que nous ne l'ayons pas encore confirmé, nos différentes données de puces ont révélé la sous-expression de *PABPC4* et *PABPC1* dans les échantillons provenant de parkinsoniens. Aux vues du rôle des protéines PABPs, cette variation d'expression laisse suggérer l'implication du NMD dans la MP. Par ailleurs, les travaux de Soreq et ses collaborateurs ont montré dans les cellules sanguines, une différence au niveau de l'induction du NMD chez les parkinsoniens avant et après un traitement par la stimulation cérébrale profonde. En plus de montrer que ce traitement induit une modification de profil d'expression des transcrits, Soreq a montré que les sujets sans stimulation avaient 319 épissages alternatifs différents par rapport aux témoins et que la stimulation entraînait une diminution du nombre de ces événements d'épissage et du nombre de ceux conduisant à des modifications des profils de NMD (c'est à dire les ARNs qui sont destinés à être dégradé).

La fonction des protéines PABPs dans la dégradation des ARNs passe également par les miARNs. PABPC1 interagit avec GW182 et Ago2 (argonautes), deux protéines qui facilitent la reconnaissance et la fixation des miARNs à leurs cibles<sup>325</sup>. L'interaction de PABPC1 avec les protéines argonautes rappelle celle de LRRK2<sup>295</sup>. En effet, le miARN miR-184 régule négativement l'expression d'*E2F1*. La surexpression de ce facteur a été montrée comme étant inductrice de la mort des neurones dopaminergiques et d'une baisse de motricité chez des drosophiles transgéniques. L'interaction de LRRK2 avec la protéine Argonaute 2 impliquée dans la reconnaissance des ARNs cibles des miARNs, favorisée par la mutation G2019S, empêche l'inhibition de *E2F1* par le miR-184<sup>223</sup>. Cette même mutation entraîne la mort des neurones dopaminergiques humains, dérivés de cellules souches, par phosphorylation de la protéine ribosomique Rps15 qui entraîne une hausse sensible de la traduction<sup>326</sup>. Une diminution de la synthèse de l' $\alpha$ -synucléine dans des lignées cellulaires de HEK293T a été montrée en réponse à une diminution du niveau de fer<sup>327</sup>. Bien que cette régulation semble se faire au niveau protéique, il est intéressant de constater la baisse des niveaux d'ARNs des

gènes *ALAS2* (impliqué dans la synthèse du fer) et de *SNCA* dans les cellules sanguines de patients qui pourrait contribuer au lien entre fer et *SNCA* par un autre mode de régulation. La synthèse de l' $\alpha$ -synucléine est inhibée par la méthylation au niveau de l'intron 1 de son gène via la méthyltransférase DNMT1. Dans le cerveau des parkinsoniens, *SNCA* est hypométhylé au niveau de l'intron 1 induisant une augmentation de sa traduction<sup>219</sup>. Sachant que la surexpression de l' $\alpha$ -synucléine est pathologique, son augmentation par perturbation de sa traduction tend à montrer l'importance de la synthèse protéique dans la MP d'autant plus que les perturbations de la traduction notamment via la mutation G2019S entraînent la mort des cellules dans un modèle de drosophile.

Dans les lignées cellulaires de carcinome gastrique humain, il a été montré que l'inhibition de PABPC1 entraînait une augmentation du niveau de protéines pro-apoptotiques BAX et le clivage de la caspase 3 ainsi qu'une diminution de la protéine anti-apoptotique Bcl-2. Cette inhibition de PABPC1 est également associée à une augmentation de l'expression de miR-34c, un miARN contribuant à l'apoptose<sup>328</sup>. Il a récemment été montré que miR-34c étaient un régulateur de l'expression de *SNCA* et que l'inhibition de ce miARN induisait une augmentation de l'expression de *SNCA*<sup>329</sup>. Ces résultats sont intéressants car même s'il n'a pas été prouvé que l'effet de PABPC4 sur l'expression miR-34c était le même que celui de PABPC1 sur ce miARN, la diminution de l'expression de *PABPC4* que nous avons observée dans les cellules sanguines de parkinsoniens pourrait conduire à une augmentation de miR-34c conduisant à son tour à la diminution de l'expression de *SNCA*. Par ailleurs, l'expression de *SNCA* est augmentée chez les sujets atteints d'ataxie spinocérébelleuse de type 2 pour qui nous n'avons pas mis en évidence de variations ni de *PABPC4* ni de *PABPC1*.

*PABPC4* est principalement exprimé dans les lymphocytes T activés (passage des lymphocytes T natifs aux lymphocytes matures). Sa surexpression est retrouvée dans les cellules tumorales et est associée à un meilleur pronostic vital<sup>330</sup>. Comme les autres protéines de la famille PABP, *PABPC4* est capable de se fixer à la queue polyA des ARNs au niveau de motif AU-rich (riches en adénosine et en uraciles) et ce principalement au niveau de queues polyA inférieures à 30 nucléotides adénosines. Cette fixation empêche les ARNm ciblés par *PABPC4* d'être dégradés par le NMD. La perte de *PABPC4* dans les cellules a été associée à une diminution de la quantité d'ARNs ciblés par *PABPC4* résultant de leur dégradation par le NMD<sup>331</sup>. Il se pourrait que la diminution de l'expression que nous avons observée chez les parkinsoniens soit associée à une diminution de certains ARNs par induction de leur dégradation via le NMD. Il est possible que *SNCA* soit un des gènes dont les ARNs seraient dégradés depuis qu'Abeliovich *et al.* ont mis en évidence l'existence de transcrits de ce gène avec des longueurs différentes de leurs extrémités 3' non traduites ; les transcrits dont l'extrémité 3'UTR est longue étant plus fortement exprimés et traduits<sup>296</sup>. Il n'est pas exclu que la prépondérance de ces

transcrits à longue extrémité 3' non traduite résulte de la dégradation des transcrits à forme courte. L'ensemble de ces données montrent aussi qu'à travers différents mécanismes, il existerait une perturbation de la traduction des ARNs dans la MP.

Si pour le moment nous n'avons pas exploré les répercussions des mutations de l'*ATXN2* sur la traduction, la localisation de l'ataxine-2 au niveau des granules de stress (corps cytoplasmiques où les ARNm sont recrutés en condition de stress pour inactiver la mort cellulaire et favoriser la survie) et son interaction avec DDX6, une hélicase impliquée dans la dégradation et la répression de la traduction, renforce l'idée de l'implication de la traduction dans la MP ; ce d'autant que la surexpression de l'ataxine-2 sauvage ou mutée (79 répétitions) dans les cellules HEK-293T et dans les SH-SY5Y où l'inhibition de la forme sauvage entraîne des perturbations de l'assemblage des granules de stress. La sous-expression de l'*ATXN2* entraîne dans ces lignées cellulaires, une augmentation de l'expression des protéines PABPs alors que la surexpression de la forme sauvage ou de la forme mutée entraîne une diminution<sup>332</sup>.

#### 4.4. L'épissage alternatif comme source de biomarqueurs potentiels de la maladie de Parkinson

Selon le NIH (National Institutes of Health), les biomarqueurs correspondent à tous les paramètres mesurables indicateurs d'un processus biologique, d'un processus pathogène ou d'une réponse à un traitement pharmacologique. Ce paramètre doit varier avec l'évolution de la maladie et sa mesure doit être sensible, spécifique et fiable. Ainsi, plusieurs types de biomarqueurs ont été proposés comme candidats potentiels : des biomarqueurs biochimiques, des biomarqueurs reposant sur des techniques d'imagerie, des biomarqueurs cliniques et des biomarqueurs génétiques qui sont les derniers à avoir été proposés et qui vont nous intéresser particulièrement.

Du fait de son rôle central dans la maladie, l' $\alpha$ -synucléine a fait l'objet de plusieurs tentatives visant à être utilisé comme biomarqueurs. Elle peut être détectée dans le liquide céphalo-rachidien, la salive, le sérum, les urines et le tractus gastro-intestinal. Un faible niveau d' $\alpha$ -synucléine a été montré dans le liquide céphalo-rachidien de patients par rapport aux témoins sans qu'il n'y ait d'association avec la sévérité de la maladie. Des essais de mesure de l' $\alpha$ -synucléine dans le sang ont été réalisés ainsi que dans le plasma mais les résultats sont contradictoires<sup>333,334</sup>; certains montrant des augmentations, d'autres des diminutions. La mesure de l' $\alpha$ -synucléine dans les biopsies gastriques est encourageante non seulement car les différentes études s'accordent à dire qu'elle est surexprimée dans ces tissus mais aussi car il s'agirait d'un marqueur précoce, le tractus gastro-intestinal étant considéré comme le point de départ de la MP. Néanmoins, le nombre d'études réalisées pour le moment reste limité et cette mesure reste invasive pour le patient. De même que

pour l' $\alpha$ -synucléine, la mesure de la protéine DJ-1 a été suggérée dans le liquide céphalo-rachidien montrant une augmentation chez les parkinsoniens qui cependant n'est pas retrouvée dans différentes cohortes<sup>335</sup>.

Plusieurs techniques d'imagerie ont été développées et permettent de visualiser la signalisation dopaminergique à l'aide de traceur et voir la dégénérescence neuronale et ce même avant l'établissement du diagnostic. Récemment des techniques d'imagerie par résonance magnétique ont été montrées comme pouvant détecter l'augmentation de fer dans la SN<sup>336</sup> ; du fait du lien entre le fer et la mort neuronale, ceci permettrait de détecter la maladie avant même que les neurones ne commencent à dégénérer.

Les symptômes non moteurs ont fait l'objet d'intérêt tout particulier dans la recherche de biomarqueurs du fait de leur apparition précoce. C'est notamment le cas l'anosmie, des troubles du sommeil ou encore des troubles intestinaux. Les troubles du sommeil sont associés à un risque augmenté de développer la MP et la démence. Les troubles de l'olfaction présents dans plus de 90% des cas de MP font aussi l'objet d'étude afin d'évaluer leur potentiel en tant que biomarqueur. Une batterie de tests incluant l'acuité olfactive et les troubles du sommeil a déjà été utilisée pour screener les patients dans plusieurs études, tests qui présentent l'avantage de ne pas être invasifs, peu coûteux mais qui ont une spécificité et une sensibilité limitées. C'est d'ailleurs le cas de l'ensemble des troubles non moteurs qui ne sont pas spécifiques de la MP.

Du fait de la diversité de la maladie, on ne peut imaginer qu'un seul marqueur reflète toute sa complexité et la stratégie qui prime dans la littérature est l'utilisation de plusieurs marqueurs combinés. Ces dernières années, les analyses du transcriptome de cellules sanguines sont apparues comme un moyen efficace d'identifier des modifications moléculaires caractéristiques d'une maladie. En effet, elles présentent un grand intérêt pour la recherche clinique en raison du caractère peu invasif et peu coûteux de la prise de sang qui est grandement rentabilisé puisqu'elle permet de déterminer une signature moléculaire du vivant du patient et ainsi de suivre la progression de la maladie et les réponses aux traitements. Aussi le fait que ces cellules reflètent plusieurs atteintes retrouvées au niveau cérébral en font des bons candidats en tant que matériel source dans la recherche de marqueurs de la maladie. Les variations que nous avons observées concernent aussi bien des gènes codant des régulateurs de l'épissage que des composés de sa machinerie effectrice tel que U3 suggérant une atteinte globale de ce mécanisme pouvant induire des variations de compositions en exons d'un grand nombre de transcrits.

Schegadeh *et al.* avait également émis cette hypothèse suite à une analyse comparée de données transcriptomiques de la littérature qui avaient montré la variation d'expression du facteur d'épissage SRRM2 (une protéine SR constituant de la machinerie effectrice de l'épissage) à la fois dans des lignées cellulaires de neuroblastome exposées à la roténone, dans la substance noire de parkinsoniens décédés et dans des lymphocytes de patients vivants. Deux transcrits de ce gène ont été identifiés pour lesquels la dérégulation a été confirmée par RT-qPCR dans la substance noire (augmentation de la forme courte et diminution de la forme longue) et dans l'amygdale (diminution de la forme longue) de parkinsoniens<sup>278</sup>. L'étude de l'épissage à l'aide de puces à exons à partir d'échantillons sanguins révélant la surexpression de la forme courte et une sous expression de la forme longue a également montré des différences de jonctions entre exons au sein de 218 gènes<sup>278</sup>. Ces dérèglements d'expression de SRRM2 ont également été retrouvés dans les PBMC de patients renforçant le lien entre l'épissage et la maladie<sup>278,337,338</sup>. Néanmoins, nous avons testés les dérèglements d'expression de transcrits de ce gène et n'avons pas trouvé de différences significatives entre malades et témoins. Il semble en effet qu'il existe d'autres transcrits que ceux décrits initialement par Shehadeh, rendant les mesures délicates.

Soreq *et al.* avait également observé, par l'utilisation de puces à exons, la variation de l'épissage de 150 gènes dans les lymphocytes de malades versus témoins, parmi lesquels ont été trouvés les gènes *DJ-1* et *SNCA*<sup>339</sup>. La perturbation de l'épissage de gènes connus comme étant liés à la maladie renforce l'importance du rôle de ce processus dans la physiopathologie. Soreq *et al.* ont également révélé la surexpression de l'ARN U1 de la sous-unité U1 du spliceosome dans les lymphocytes, la substance noire et l'amygdale de patients parkinsoniens<sup>337</sup>. Certains traitements des symptômes cliniques de la MP ont également un effet sur le profil des transcrits exprimés dans le cerveau des malade, ce qui a été montré par l'étude de Soreq *et al.* révélant que la stimulation cérébrale profonde de parkinsoniens peut induire une modification de la signature des transcrits exprimés dans le sang de ces mêmes sujets. Ces modifications réversibles impliquent des gènes tels que *SNCA* et *PARK7* mais également des gènes impliqués dans les processus associés à la MP tels que l'inflammation<sup>298,340</sup>. Bien que les gènes ne soient pas les mêmes entre les différentes études, l'observation de variations d'expression de facteurs liés à l'épissage est un argument en faveur de l'existence de modifications de ce processus chez les parkinsoniens et de leur rôle dans la physiopathologie de la maladie.

## 5. Conclusion

Ce travail de thèse avait pour but de répondre à plusieurs objectifs visant dans un premier temps à déterminer si les similitudes phénotypiques entre les parkinsoniens sporadiques et les parkinsoniens porteurs de la mutation G2019S de *LRRK2* reposaient sur des perturbations moléculaires communes. Ces sujets porteurs de la mutation de *LRRK2* ayant initialement montraient des variations d'expression de gènes liés au métabolisme des ARNs, cette comparaison a permis de révéler que ces modifications pouvaient être retrouvées chez un grand nombre de cas. En effet, la mise en évidence de variations d'expression génique associées à la voie de signalisation EIF2 est commune aux sujets porteurs de la mutation G2019S de *LRRK2* et aux cas sporadiques. Aux vues du lien entre cette voie de signalisation et la traduction ainsi que l'épissage, cette découverte a fourni un premier argument en faveur de l'existence de perturbations de ces processus chez un grand nombre de cas. L'analyse du transcriptome de sujets parkinsoniens porteurs des mutations de *ATXN2* a appuyé ces données en montrant une fois de plus des variations d'expression de gènes liés au métabolisme des ARNs. L'étude du transcriptome par l'utilisation du RNAseq a définitivement marqué l'implication du métabolisme des ARNs et en particulier de l'épissage dans la MP.

Bien que nous soyons conscients des limites de nos analyses en partie du fait du faible nombre de sujets étudiés, l'analyse des voies et des processus biologiques liés aux perturbations mises en évidence par ces différentes études convergent toutes vers l'identification de perturbations de la réponse inflammatoire, la réponse immunitaire, la prolifération, la mort cellulaire, l'endocytose, la mitochondrie et de l'ubiquitination. Ainsi malgré le manque de puissance de notre étude, la mise en évidence par ces différentes analyses de voies et processus biologiques similaires, qui plus est, validés par les données de la littérature, nous portent à croire que les variations que nous avons observées ne sont pas du fait du hasard. Les validations expérimentales des jonctions nouvelles des gènes *RBM5* et *CCLN1* sont également la preuve de la pertinence des résultats de notre étude qui mérite d'être poursuivie. Cette démarche avait également pour but d'évaluer si les variations du métabolisme des ARNs pouvaient constituer un mécanisme physiopathologique de la MP survenant dès les premiers stades et si elles pouvaient permettre de détecter et de traiter la maladie plus précocément. Nous avons montré par l'analyse des sujets *de novo* que les variations quantitatives et qualitatives de gènes liés à l'épissage étaient déjà présentes chez ces sujets.

De façon intéressante, chez ces mêmes sujets, les variations quantitatives d'expression de transcrits et les modifications de jonctions entre exons affectent principalement des gènes liés à la réponse immunitaire. Dans ce contexte actuel où plusieurs études tendent à montrer le rôle de la

réponse immunitaire dans la génèse de la maladie, le futur de notre étude pourra déterminer si les changements précoces de l'immunité résultent de ces défauts d'épissage.

# **ANNEXES**



## 6. Annexes

### **Annexe 1: Critères diagnostiques UKPDSBB de la MP (United Kingdom Parkinson's Disease Society Brain Bank)**

**1ère étape :** Diagnostic d'un syndrome parkinsonien :

- bradykinésie (lenteur d'initiation des mouvements volontaires avec réduction progressive de la vitesse et de l'amplitude d'actions répétées)
- et au moins un des signes suivants :
  - rigidité musculaire
  - tremblement de repos à 4 à 6 Hz
  - instabilité posturale non expliquée par un trouble visuel, vestibulaire, cérébelleux ou proprioceptif

**2ème étape :** Absence d'un des critères d'exclusion suivants :

- histoire d'accidents vasculaires répétés avec progression par à-coups des symptômes parkinsoniens
- antécédent de traumatismes crâniens répétés
- antécédent d'encéphalite
- crises oculogyres
- traitement neuroleptique précédant l'installation des symptômes
- antécédents familiaux de plusieurs cas de syndrome parkinsonien
- rémission prolongée du syndrome parkinsonien
- signes strictement unilatéraux après 3 ans d'évolution
- paralysie supranucléaire du regard
- syndrome cérébelleux

- dysautonomie sévère et précoce
- démence sévère et précoce avec troubles de la mémoire, du langage et des praxies
- signe de Babinski
- présence d'une tumeur cérébrale ou d'une hydrocéphalie communicante au scanner
- réponse négative à des doses élevées de Lévodopa (en l'absence de malabsorption)
- exposition au MPTP.

**3ème étape** : Critères positifs en faveur d'une maladie de Parkinson :

(au moins trois de ces critères sont nécessaires pour poser le diagnostic d'une MP définie)

- début unilatéral
- présence d'un tremblement de repos
- évolution progressive
- asymétrie persistante prédominant sur le côté initialement atteint,
- amélioration significative du syndrome parkinsonien par la levodopa (plus de 70 %)
- mouvements choréïques intenses induits par la levodopa
- sensibilité à la levodopa pour une durée supérieure ou égale à 5 ans
- durée de l'évolution clinique supérieure ou égale à 10 ans

## **Annexe 2: Échelle de Hoehn et Yahr**

Elle évalue la sévérité globale du syndrome parkinsonien selon 5 stades :

Stade 1 = maladie unilatérale

Stade 1.5 = maladie unilatérale avec atteinte axiale

Stade 2 = maladie bilatérale sans trouble de l'équilibre

Stade 2.5 = maladie bilatérale légère avec retentissement lors du test de la poussée

Stade 3 = maladie bilatérale légère à modérée, certaine instabilité posturale, physiquement autonome

Stade 4 = handicap sévère, toujours capable de marcher ou de se tenir debout sans aide

Stade 5 = malade en chaise roulante ou alité, n'est plus autonome

**Annexe 3 : Échelle d'évaluation unifiée pour la maladie de Parkinson (UPDRS, Unified Parkinson Disease Rating Scale)**

I- Etat mental, comportemental et thymique

1- Affaiblissement intellectuel

0 = absent.

1 = léger. Manque de mémoire habituel avec souvenir partiel des événements.

2 = perte mnésique modérée, désorientation, difficultés modérées à faire face à des problèmes complexes. Atteinte légère mais indiscutable des capacités fonctionnelles avec besoin d'une incitation occasionnelle de l'entourage.

3 = déficit mnésique grave avec désorientation dans le temps et souvent dans l'espace. Handicap grave face aux problèmes.

4 = perte mnésique sévère avec uniquement préservation de sa propre orientation. Incapacité de porter des jugements et de résoudre des problèmes, demande d'aide pour les soins personnels, ne peut être laissé seul.

2- Troubles de la pensée

0 = aucun.

1 = rêves animés.

2 = hallucinations bénignes critiquées.

3 = hallucinations occasionnelles ou fréquentes ou idées délirantes non critiquées.

4 = hallucinations continuelles, psychose expansive, incapacité de prendre soin de soi.

3- Dépression

0 = absente.

1 = périodes de tristesse ou sentiment de culpabilité excessif ne persistant pas plusieurs jours.

2 = dépression durable (une semaine ou plus).

3 = dépression durable avec symptômes végétatifs (insomnie, anorexie, perte de poids).

4 = dépression durable avec troubles végétatifs, pensées ou intentions suicidaires.

#### 4- Motivation – Initiative

0 = normale.

1 = moins franche que d'habitude, plus passif.

2 = perte d'initiative et désintérêt pour certaines activités non routinières.

3 = perte d'initiative ou désintérêt dans les activités routinières.

4 = absence d'initiative, perte totale d'intérêt.

#### II- Activités de la vie quotidienne (à déterminer en période ON et en période OFF)

##### 5- Parole

0 = normale.

1 = légèrement perturbée, pas de difficulté à être compris.

2 = modérément perturbée. On doit lui demander occasionnellement de répéter.

3 = gravement perturbée. On doit lui demander fréquemment de répéter.

4 = incompréhensible la plupart du temps.

##### 6- Salivation

0 = normale.

1 = légère, mais excès habituel de salive dans la bouche, peut baver la nuit.

2 = hypersialorrhée modérée, peut baver un peu.

3 = hypersialorrhée nette avec un peu de bave.

4 = écoulement habituel de bave nécessitant en permanence un mouchoir.

166

#### 7- Déglutition

0 = normale.

1 = s'étrangle rarement.

2 = s'étrangle occasionnellement.

3 = nécessite une alimentation semi-liquide.

4 = nécessite une alimentation par sonde gastrique ou une gastrectomie.

#### 8- Ecriture

0 = normale.

1 = légèrement ralentie ou micrographique.

2 = nettement ralentie ou micrographique, tous les mots sont lisibles.

3 = gravement perturbée, tous les mots ne sont pas lisibles.

4 = la majorité des mots est illisible.

#### 9- S'alimenter et manipuler les couverts

0 = normal.

1 = un peu lent et maladroit mais n'a pas besoin d'être aidé.

2 = peut se débrouiller seul pour la plupart des aliments mais lent et maladroit.

3 = a besoin d'une aide pour les repas mais peut encore s'alimenter lentement.

4 = on doit lui donner à manger.

#### 10- Habillage

0 = normal.

1 = un peu lent mais ne doit pas être aidé.

2 = aide occasionnelle pour boutonner ou enfiler une manche.

3 = a besoin d'être très aidé mais peut faire certaines choses seul.

4 = totalement dépendant.

#### 11- Hygiène

0 = normale.

1 = un peu lent, mais n'a pas besoin d'aide.

2 = nécessite une aide pour la douche ou le bain, très lent dans les soins hygiéniques.

3 = nécessite une aide pour se laver, se brosser les dents, se coiffer, se baigner.

4 = sonde urinaire ou autres aides mécaniques.

#### 12- Se retourner dans le lit, arranger les draps et couvertures

0 = normal.

1 = un peu lent et maladroit mais n'a pas besoin d'être aidé.

2 = peut se retourner seul ou arranger les draps mais avec une grande difficulté.

3 = peut commencer le geste mais n'arrive pas à se retourner ou arranger les draps.

4 = dépendant.

#### 13- Chute non liée au piétinement

0 = aucune.

1 = chutes rares.

2 = chutes occasionnelles mais moins d'une fois par jour.

3 = en moyenne une chute par jour.

4 = chutes pluriquotidiennes.

14- Piétinement lors de la marche

0 = aucun.

167

1 = rare, peut avoir une hésitation au départ.

2 = piétinement occasionnel lors de la marche.

3 = piétinement fréquent entraînant occasionnellement des chutes.

4 = chutes fréquentes dues au piétinement.

15- Marche

0 = normale.

1 = difficultés légères, peut balancer les bras, traîne les pieds.

2 = difficultés modérées mais ne demande que peu ou pas d'aide.

3 = difficultés importantes nécessitant une aide.

4 = ne peut marcher du tout même avec aide.

16- Tremblement

0 = absent.

1 = léger et rarement présent.

2 = modéré, gênant le patient.



3 = important, gênant certaines activités.

4 = marqué, gênant la plupart des activités.

#### 17- Troubles subjectifs sensitifs liés au syndrome parkinsonien

0 = aucun.

1 = occasionnellement, engourdissements, picotements, douleurs légères.

2 = engourdissements, picotements, douleurs fréquentes, pas gênant.

3 = sensations douloureuses fréquentes.

4 = douleurs très vives.

#### III- Examen moteur

##### 18- Parole

0 = normale.

1 = légère perte d'expression, de la diction et/ou du volume vocal.

2 = voix monotone, bredouillée mais compréhensible, altération modérée.

3 = altération marquée, difficile à comprendre.

4 = incompréhensible.

##### 19- Expression faciale

0 = normale.

1 = hypomimie légère, semble avoir un visage normalement impassible.

2 = diminution légère mais franchement anormale de l'expression faciale.

3 = hypomimie modérée, lèvres souvent entrouvertes.

4 = masque facial avec perte importante de l'expression faciale, lèvres entrouvertes.

## 20- Tremblement de repos

0 = absent.

1 = léger et rarement présent.

2 = tremblement de faible amplitude mais persistant. Ou amplitude modérée mais intermittent.

3 = tremblement modéré en amplitude et présent la plupart du temps.

4 = tremblement d'amplitude marquée et présent la plupart du temps.

## 21- Tremblement d'action ou tremblement postural des mains

0 = absent.

1 = léger : présent lors de l'action.

168

2 = modéré en amplitude, présent lors de l'action.

3 = modéré en amplitude, tant lors du maintien postural que lors de l'action.

4 = amplitude marquée : gêne l'alimentation.

## 22- Rigidité

0 = absent.

1 = minime ou apparaissant lors des manœuvres de sensibilisation.

2 = légère à modérée.

3 = marquée, mais la plupart des mouvements peuvent être effectués aisément.

4 = sévère, les mouvements sont effectués difficilement.

## 23- Tapotement des doigts

0 = normal.

1 = ralentissement léger et/ou réduction d'amplitude.

2 = modérément perturbé, se fatigue rapidement, occasionnels arrêts du mouvement.

3 = sévèrement perturbé, hésitation au début du mouvement ou arrêt du mouvement.

4 = peut à peine effectuer la tâche.

#### 24- Mouvements des mains

0 = normal.

1 = ralentissement léger et/ou réduction d'amplitude.

2 = modérément perturbé, se fatigue rapidement, occasionnels arrêts du mouvement.

3 = sévèrement perturbé, hésitation au début du mouvement ou arrêt du mouvement.

4 = peut à peine effectuer la tâche.

#### 25- Mouvements alternatifs rapides

0 = normal.

1 = ralentissement léger et/ou réduction d'amplitude.

2 = modérément perturbé, se fatigue rapidement, occasionnels arrêts du mouvement.

3 = sévèrement perturbé, hésitation au début du mouvement ou arrêt du mouvement.

4 = peut à peine effectuer la tâche.

#### 26- Agilité de la jambe

0 = normal.

1 = ralentissement léger et/ou réduction d'amplitude.

2 = modérément perturbé, se fatigue rapidement, occasionnels arrêts du mouvement.

3 = sévèrement perturbé, hésitation au début du mouvement ou arrêt en cours de mouvement.

4 = peut à peine effectuer la tâche.

#### 27- Se lever d'une chaise

0 = normal.

1 = lentement ou a besoin de plus d'un essai.

2 = se pousse sur les bras du siège.

3 = tend à tomber en arrière et doit essayer plus d'une fois mais peut se lever sans aide.

4 = incapable de se lever sans aide.

#### 28- Posture

0 = normalement droite.

1 = pas tout à fait droite, posture légèrement fléchie, normale pour une personne plus âgée.

2 = posture modérément fléchie, nettement anormale : peut être légèrement penché d'un côté.

3 = flexion sévèrement fléchie avec cyphose : peut être modérément penchée d'un côté.

169

4 = flexion marquée avec posture très anormale.

#### 29- Stabilité posturale

0 = normale.

1 = rétroimpulsion mais l'équilibre est rétabli sans aide.

2 = absence de réponse posturale : peut tomber s'il n'est pas retenu par l'examineur.

3 = très instable, tend à perdre l'équilibre spontanément.

4 = incapable de se tenir debout sans aide.

#### 30- Démarche

0 = normale.

1 = marche lentement, peut trainer les pieds et faire des petits pas, mais sans festination.

2 = marche avec difficultés ; peut avoir un peu de festination, des petits pas.

3 = perturbations sévères de la marche, nécessitant une aide.

4 = ne peut marcher du tout, même avec aide.

### 31- Bradykinésie corporelle

0 = aucune.

1 = lenteur minime, donnant aux mouvements un caractère délibéré. Réduction d'amplitude.

2 = lenteur et pauvreté du mouvement qui est nettement anormal.

3 = lenteur modérée, pauvreté et petite amplitude du mouvement.

4 = lenteur marquée, pauvreté et petite amplitude du mouvement.

## IV- Complications du traitement (au cours de la dernière semaine)

### A- Dyskinésies

#### 32- Durée

0 = aucune.

1 = 1 à 25% de la journée.

2 = 26 à 50% de la journée.

3 = 51 à 75% de la journée.

4 = 76 à 100% de la journée.

#### 33- Incapacité

0 = aucune.

1 = légère.

2 = modérée.

3 = sévère.

4 = complète.

34- Douleur

0 = aucune.

1 = légère.

2 = modérée.

3 = sévère.

4 = complète.

35- Présence d'une dystonie matinale précoce

0 = non.

1 = oui.

170

B- Fluctuations motrices

36- Y a-t-il des périodes off dont on peut prédire le moment après une prise médicamenteuse ?

0 = non.

1 = oui.

37- Y a-t-il des périodes off dont on ne peut pas prédire le moment après une prise médicamenteuse ?

0 = non.

1 = oui.

38- Est-ce que certaines des périodes off viennent soudainement ?

0 = non.

1 = oui.

39- Quelle est la proportion de la journée durant laquelle le patient est en situation off ?

0 = aucune.

1 = 1 à 25% de la journée.

2 = 26 à 50% de la journée.

3 = 51 à 75% de la journée.

4 = 76 à 100% de la journée.

C- Autres complications

40- Le patient a-t-il de l'anorexie, des nausées, des vomissements ?

0 = non.

1 = oui.

41- Le patient a-t-il une hypotension orthostatique symptomatique ?

0 = non.

1 = oui.

V- Stade de Hoehn et Yahr

- Stade 1 = maladie unilatérale.

- Stade 1,5 = maladie unilatérale avec atteinte axiale.

- Stade 2 = maladie bilatérale sans trouble de l'équilibre.

- Stade 2,5 = maladie bilatérale légère avec rétablissement lors du test de la poussée.
- Stade 3 = maladie bilatérale légère à modérée, instabilité posturale, physiquement autonome.
- Stade 4 = handicap sévère, toujours capable de marcher ou de se tenir debout sans aide.
- Stade 5 = malade en chaise roulante ou alité, n'est plus autonome.

#### VI- Echelle de Schwab et England

- 100% = totalement indépendant.
- 90% = indépendant mais plus lent.
- 80% = indépendant mais conscient de sa lenteur.
- 70% = pas tout à fait indépendant, 3 à 4 fois plus lent.
- 60% = partiellement dépendant.
- 50% = aidé dans 50% des activités.
- 40% = très dépendant.
- 30% = peu d'activités effectuées seul.
- 20% = ne fait rien seul, aide légèrement.
- 10% = alité, totalement dépendant.
- 0% = alité, troubles végétatifs.



**Annexe 4 : Critères de diagnostic clinique du NINDS (National Institute of Neurological Disorders and Stroke)**

Critères A : caractéristiques de la MP

- Tremblement de repos
- Bradykinésie
- Rigidité
- Début asymétrique

Critères B : suggérant un diagnostic différentiel

- Caractéristiques inhabituelles en début d'évolution
- Instabilité posturale franche dans les 3 premières années
- Freezing dans les 3 premières années
- Hallucinations non attribuées à des médicaments dans les 3 premières années
- Démence inaugurale ou dès la première année
- Paralyse supranucléaire du regard ou ralentissement des saccades verticales
- Dysautonomie sévère et symptomatique non liée aux médications
- Mise en évidence d'une autre étiologie (comme des lésions cérébrales de localisation compatible ou utilisation de neuroleptiques dans les 6 derniers mois)

Critères diagnostiques proposés :

- Maladie de Parkinson POSSIBLE

Au moins 2 des 4 critères A, dont au moins 1 est tremblement ou bradykinésie

+ Aucun des critères B OU symptomatologie évoluant depuis moins de 3 ans sans critère B

+ Réponse franche et durable à la levodopa ou aux agonistes dopaminergiques OU absence de traitement dopaminergique adéquat

☒ Maladie de Parkinson PROBABLE

Au moins 3 des 4 critères A

+ Aucun des critères B après 3 ans d'évolution

+ Réponse franche et durable à la levodopa ou aux agonistes dopaminergiques

☒ Maladie de Parkinson DEFINIE

Tous les critères de MP possible

+ Confirmation histopathologique en autopsie

# REFERENCES

## 7. Références

1. Pringsheim T, Jette N, Frolkis A, Steeves TD. The prevalence of Parkinson's disease: A systematic review and meta-analysis. *Mov. Disord.* 2014;29(13):1583–90.
2. Van Den Eeden SK, Tanner CM, Bernstein AL, Fross RD, Leimpeter A, Bloch DA, *et al.* Incidence of Parkinson's disease: variation by age, gender, and race/ethnicity. *Am. J. Epidemiol.* 2003;157(11):1015-22.
3. Mayeux R, Marder K, Cote LJ, Denaro J, Hemenegildo N, Mejia H, *et al.* The frequency of idiopathic Parkinson's disease by age, ethnic group, and sex in northern Manhattan, 1988-1993. *Am. J. Epidemiol.* 1995;142(8):820-7.
4. Wright Willis A, Evanoff BA, Lian M, Criswell SR, Racette BA. Geographic and Ethnic Variation in Parkinson Disease: A Population-Based Study of US Medicare Beneficiaries. *Neuroepidemiology.* 2010;34(3):143-51.
5. de Rijk MC, Breteler MM, Graveland GA, Ott A, Grobbee DE, van der Meché FG *et al.* Prevalence of Parkinson's disease in the elderly: the Rotterdam Study. *Neurology.* 1995;45(12):2143-6.
6. Dawson TM, Dawson VL. Rare genetic mutations shed light on the pathogenesis of Parkinson disease. *J. Clin. Invest.* 2003;111(2):145-51.
7. Dorsey ER, Constantinescu R, Thompson JP, Biglan KM, Hollowy RG, Kieburtz K, *et al.* Projected number of people with Parkinson disease in the most populous nations, 2005 through 2030. *Neurology.* 2007;68(5):384-6.
8. Parkinson J. An Essay on the Shaking Palsy. *J. Neuropsychiatr.* 2002;14(2):223-36.
9. Xia R, Mao ZH. Progression of motor symptoms in Parkinson's disease. *Neurosci. Bull.* 2012;28(1):39-48.
10. Langston JW. The parkinson's complex: parkinsonism is just the tip of the iceberg. *Ann. Neurol.* 2006;59(4):591-6.
11. Jankovic J. Parkinson's disease: clinical features and diagnosis. *J. Neurol. Neurosurg. Psychiatry.* 2008;79(4):368–76.
12. Collins L, Cummins G, Barker RA. Parkinson's disease : diagnosis and current management. *Prescriber.* 2015;26(5):16-23.
13. Tolosa E, Gaig C, Santamaria J, Compta Y. Diagnosis and the premotor phase of Parkinson disease. *Neurology.* 2009;72(7 Suppl):S12-20.
14. C., T. Contribution à l'étude de l'anatomie pathologique du locus niger de Sœmmering avec quelques déductions relatives à la pathogénie des troubles du tonus musculaire et de la maladie de Parkinson. Paris: Jouve; 1919.
15. Levey AI, Hersch SM, Rye DB, Sunahara RK, Niznik HB, Kitt CA, *et al.* Localization of D1 and D2 dopamine receptors in brain with subtype-specific antibodies. *Proc. Natl. Acad. Sci. U. S. A.* 1993;90(19):8861-5.
16. Kish SJ, Shannak K, Hornykiewicz O. Uneven pattern of dopamine loss in the striatum of patients with idiopathic Parkinson's disease. Pathophysiologic and clinical implications. *N. Engl. J. Med.* 1988;318(14):876-80.
17. Parent M, Parent A. Substantia nigra and Parkinson's disease : a brief history of their long and intimate relationship. *Can. J. Neurol. Sci.* 2010.37(3):313-9.
18. Brooks DJ. The Early Diagnosis of Parkinson's Disease. *Ann Neurol.* 1998;44(3 Suppl 1):S10–8.

19. Kastner A, Hirsch EC, Lejeune O, Javoy-Agid F, Rascol O, Agid Y. Is the vulnerability of neurons in the substantia nigra of patients with Parkinson's disease related to their neuromelanin content? *J. Neurochem.* 1992;59(3):1080-9.
20. Ma SY, Røyttä M, Rinne JO, Collan Y, Rinne UK. Correlation between neuromorphometry in the substantia nigra and clinical features in Parkinson's disease using disector counts. *J. Neurol. Sci.* 1997;151(1):83-7.
21. Hirsch EC, Graybiel AM, Agid Y. Selective vulnerability of pigmented dopaminergic neurons in Parkinson's disease. *Acta Neurol Scand Suppl.* 1989;126:19–22.
22. Spillantini MG, Schmidt ML, Lee VM, Trojanowski JQ, Jakes R, Goedert M. Alpha-synuclein in Lewy bodies. *Nature.* 1997;388(6645):839-40.
23. Spillantini MG, Crowther RA, Jakes R, Hasegawa M, Goedert M. alpha-synuclein in filamentous inclusions of Lewy bodies from Parkinson's disease and dementia with Lewy bodies. *Proc. Natl. Acad. Sci.* 1998;95(1):6469-73.
24. Hughes AJ, Daniel SE, Kilford L, Lees AJ. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. *J. Neurol. Neurosurg. Psychiatry.* 1992;55(3):181-4.
25. Calne DB, Snow BJ, Lee C. Criteria for Diagnosing Parkinson's Disease. 1992;32 Suppl:S125-7.
26. Hoehn MM, Yahr MD. Parkinsonism:onset, progression, and mortality. *Neurology.* 2001;57(10 Suppl 3):S11-26.
27. Bhidayasiri, R. & Tarsy, D. Movement Disorders: A Video Atlas. 4–6 (2012). doi:10.1007/978-1-60327-426-5
28. The Unified Parkinson's Disease Rating Scale ( UPDRS ): Status and Recommendations. *Mov. Disord.* 2003.18/738-50.
29. Calhoun CL. Treatment of Parkinson's Disease with Levodopa. *J. Natl. Med. Assoc.* 1972;64(6):516-passim.
30. Ossig C, Reichmann H. Treatment of Parkinson's disease in the advanced stage. *J. Neural Transm.* 2013;120(4):523-9.
31. Boyle A, Ondo W. Role of Apomorphine in the Treatment of Parkinson's Disease. *CNS Drugs* 2015;29(2):83-9.
32. Schneeberger A, Mandler M, Mattner F, Schmidt W. Vaccination for Parkinson's disease. *Parkinsonism Relat. Disord.* 2012;18 Suppl 1:S11-3.
33. Masliah E, Rockenstein E, Adame A, Alford M, Crews L, Hashimoto M, *et al.* Effects of alpha-synuclein immunization in a mouse model of Parkinson's disease. *Neuron.* 2005;46(6):857-68.
34. Mandler M, Vamera E, Rockenstein E, Weninger H, Patrick C, Adame A, *et al.* Next-generation active immunization approach for synucleinopathies: Implications for Parkinson's disease clinical trials. *Acta Neuropathol.* 2014;127(6):861-79.
35. Olanow CW, Rascol O, Hauser R, Feigin PD, Jankovic J, Lang A, *et al.* A double-blind, delayed-start trial of rasagiline in Parkinson's disease. *N. Engl. J. Med.* 2009;361(13):1268-78.
36. Braak H, Del Tredici K, Rüb U, de Vos R, Jansen Steur EN, Braak E. Staging of brain pathology related to sporadic Parkinson's disease. *Neurobiol. Aging.* 2003;24(2):197-211.
37. Ojo OO, Okubadejo NU, Ojini FI, Danesi MA. Frequency of cognitive impairment and depression in Parkinson's disease: A preliminary case-control study. *Niger. Med. J.* 2012;53(2):65-70.
38. Kirsch-Darrow L, Marsiske M, Okun MS, Bauer R, Bowers D. Apathy and Depression: Separate Factors in Parkinson's Disease. *J. Int. Neuropsychol. Soc.* 2011;17(6):1058-66.

39. Mouret J. Differences in sleep in patients with Parkinson's disease. *Electroencephalogr. Clin. neurophysiol.* 1975;38:56°/653-7.
40. Roychowdhury S, Forsyth DR. Sleep disturbance in Parkinson disease. *J. Clin. Gerontol. Geriatr.* 2012;3(2):53-61.
41. Lang AE. A critical appraisal of the premotor symptoms of Parkinson's disease: Potential usefulness in early diagnosis and design of neuroprotective trials. *Mov. Disord.* 2011;26(5):775-83.
42. Braak H, Del Tredici K, Gai WP, Braak E. Alpha-synuclein is not a requisite component of synaptic boutons in the adult human central nervous system. *J. Chem. Neuroanat.* 2000;20(3-4):245-52.
43. Driver-Dunckley E, Adler CH, Hentz JG, Dugger BN, Shil HA, Caviness JN, *et al.* Olfactory dysfunction in incidental Lewy body disease and Parkinson's disease. *Parkinsonism Relat. Disord.* 2014;20(11):1260-2.
44. Paschen L, Schmidt N, Wolff S, Cnyrim C, van Eimeren T, Zeuner KE, *et al.* The olfactory bulb volume in patients with idiopathic Parkinson's disease. *Eur. J. Neurol.* 2015;22(7):1068-73.
45. Paillusson S, Lebouvier T, Pouclet H, Coron E, Bruley des Varannes S, Damier P, *et al.* Enteric nervous system and Parkinson's disease. *Rev. Med Interne.* 2012;33(6):335-8.
46. Tufekci KU, Meuwissen R, Genc S, Genc K. Inflammation in parkinson's disease. *Adv Protein Chem Struct Biol.* 2012;88:69-132.
47. Doorn KJ, Lucassen PJ, Boddeke HW, Prins M, Berendse HW, Drukarch B, *et al.* Emerging roles of microglial activation and non-motor symptoms in Parkinson's disease. *Prog Neurobiol.* 2012;98(2):222-38.
48. Barnum CJ, Tansey MG. Neuroinflammation and non-motor symptoms: The dark passenger of parkinson's disease? *Curr. Neurol. Neurosci. Rep.* 2012;12(4):350-8.
49. Guilarte TR. Manganese and Parkinson's Disease: A Critical Review and New Findings. *Cien. Saude Colet.* 2011;16(11):4549-66.
50. Guilarte TR. Manganese neurotoxicity: New perspectives from behavioral, neuroimaging, and neuropathological studies in humans and non-human primates. *Front. Aging Neurosci.* 2013;5:23.
51. Kim JW, Kim Y, Cheong HK, Ito K. Manganese induced parkinsonism : A case report. *J. Korean Med. Sci.* 198;13(4):437-9.
52. Guilarte TR, Burton NC, McGlothlan JL, Verina T, Zhou Y, Alexander M, *et al.* Impairment of nigrostriatal dopamine neurotransmission by manganese is mediated by pre-synaptic mechanism(s): Implications to manganese-induced parkinsonism. *J. Neurochem.* 2008;107(5):1236-1247 (2008).
53. Van Meter JR. Therapy of Parkinson's Disease. *Calif. Med.* 1950;73(4):322-24.
54. Langston JW, Ballard P, Tetrud JW, Irwin I. Chronic Parkinsonism in humans due to a product of meperidine-analog synthesis. *Science.* 193;219(4587):979-80.
55. Langston JW, Langston EB, Irwin I. MPTP-induced parkinsonism in human and non-human primates--clinical and experimental aspects. *Acta Neurol Scand Suppl.* 1984;100:49-54.
56. Moon HE, Paek SH. Mitochondrial Dysfunction in Parkinson's Disease. *Exp. Neurobiol.* 2015;24(2):103-16.
57. Keane PC, Kurzawa M, Blain PG, Morris CM. Mitochondrial dysfunction in Parkinson's disease. *Parkinsons. Dis.* 2011;2011:716871.
58. Bagchi M, Zafra-Stone S, Bagchi D, Patel S. Oxidative stress and neurodegeneration. *Vet.*

- Toxicol.* 2007;1201.
59. Winklhofer KF, Haass C. Mitochondrial dysfunction in Parkinson's disease. *Biochim. Biophys. Acta.* 2010;1802(1):29-44.
  60. Mizuno Y, Ohta S, Tanaka M, Takamiya S, Suzui K, Sato T, *et al.* Deficiencies in complex I subunits of the respiratory chain in Parkinson's disease. *Biochem. Biophys. Res. Commun.* 1989;163(3):1450-5.
  61. Nakamura K, Bindokas VP, Marks JD, Wright DA, Frim DM, Miller RJ, *et al.* The selective toxicity of 1-methyl-4-phenylpyridinium to dopaminergic neurons: the role of mitochondrial complex I and reactive oxygen species revisited. *Mol. Pharmacol.* 2000;58(2):271-8.
  62. Spina MB, Cohen G. Dopamine turnover and glutathione oxidation: implications for Parkinson disease. *Proc. Natl. Acad. Sci. U. S. A.* 1989;86(4):1398-400.
  63. Manuscript, A. GENETIC SUSCEPTIBILITY IN PARKINSON'S DISEASE. *Changes* 29, 997–1003 (2012).
  64. Hsieh YC, Mounsey RB, Teismann P. MPP(+)-induced toxicity in the presence of dopamine is mediated by COX-2 through oxidative stress. *Naunyn. Schmiedeberg's. Arch. Pharmacol.* 2011;384(2):157-67.
  65. Jana S, Sinha M, Chanda D, Roy T, Banerjee K, Munshi S, *et al.* Mitochondrial dysfunction mediated by quinone oxidation products of dopamine: Implications in dopamine cytotoxicity and pathogenesis of Parkinson's disease. *Biochim. Biophys. Acta.* 2011;1812(6):663-73.
  66. Belluzzi E, Bisaglia M, Lazzarini E, Tabares LC, Beltramini M, Bubacco L. Human SOD2 modification by dopamine quinones affects enzymatic activity by promoting its aggregation: possible implications for Parkinson's disease. *PLoS One.* 2012;7(6):e38026.
  67. Pezzella A, D'Ischia M, Napolitano A, Misuraca G, Prota G. Iron-mediated generation of the neurotoxin 6-hydroxydopamine quinone by reaction of fatty acid hydroperoxides with dopamine: A possible contributory mechanism for neuronal degeneration in parkinson's disease. *J. Med. Chem.* 1997;40(14):2211-6.
  68. Meng F, Yao D, Shi Y, Kabakoff J, Wu W, Reicher J, *et al.* Oxidation of the cysteine-rich regions of parkin perturbs its E3 ligase activity and contributes to protein aggregation. *Mol. Neurodegener.* 2011;6:34.
  69. Playfer JR. Parkinson's disease. *Postgrad. Med. J.* 1997;73(859):257-64.
  70. Vieregge P, Hagenah J, Heberlein I, Klein C, Ludin HP. Parkinson's disease in twins:a follow-up study. *Neurology.* 1999;53(3):56-72.
  71. Ward CD, Duvoisin RC, Ince SE, Nutt JD, Eldridge R, Calne DB. Parkinson's disease in 65 pairs of twins and in a set of quadruplets. *Neurology.* 1983;3(7):815-24.
  72. Schapira AH. Nuclear and mitochondrial genetics in Parkinson's disease. *J. Med. Genet.* 1995;32(6):411-4.
  73. Nussbaum RL, Polymeropoulos MH. Genetics of Parkinson's disease. *Hum. Mol. Genet.* 1997;6(10):1687-91.
  74. Polymeropoulos MH, Lavedan C, Leroy E, Ide SE, Dehejia A, Dutra A, *et al.* Mutation in the alpha-Synuclein Gene Identified in Families with Parkinson's Disease. *Science.* 1997;276(5321):2045-7.
  75. Klein, C. & Westenberger, A. Genetics of Parkinson's disease. *Cold Spring Harb. Perspect. Med.* 2, a008888 (2012).
  76. Panas, M., Karadimas, C., Avramopoulos, D. & Vassilopoulos, D. Lack of mutation G209A in French patients with familial and sporadic Parkinson's disease.

77. Kiely AP, Asi YT, Kara E, Limousin P, Ling H, Lewis P, *et al.*  $\alpha$ -Synucleinopathy associated with G51D SNCA mutation: a link between Parkinson's disease and multiple system atrophy? *Acta Neuropathol.* 2013;125(5):753-69.
78. Chartier-Harlin MC, Kachergus J, Roumier C, Mouroux V, Douay X, Lincoln S, *et al.*  $\alpha$ -synuclein locus duplication as a cause of familial Parkinson's disease. *Lancet.* 2004;364(9440):1167-9.
79. Singleton AB, Farrer M, Johnson J, Singleton A, Hague S, Kachergus J, *et al.* alpha-Synuclein locus triplication causes Parkinson's disease. *Science.* 2003;302(5646):841.
80. Vaughan JR, Farrer MJ, Wszolek ZK, Gasser T, Durr A, Agid Y, *et al.* Sequencing of the alpha-synuclein gene in a large series of cases of familial Parkinson's disease fails to reveal any further mutations. The European Consortium on Genetic Susceptibility in Parkinson's Disease (GSPD). *Hum. Mol. Genet.* 1988;7(4):751-3.
81. Corti O, Lesage S, Brice A. What Genetics Tells us About the Causes and Mechanisms of Parkinson's Disease. *Physiol. Rev.* 2011;91(4):1161-218.
82. Uéda K, Saitoh T, Mori H. Tissue-dependent alternative splicing of mRNA for NACP, the precursor of non-A beta component of Alzheimer's disease amyloid. *Biochem. Biophys. Res. Commun.* 1994;205(2):1366-72.
83. Maroteaux L, Campanelli JT, Scheller RH. Synuclein : A Neuron-Specific Presynaptic Nerve Terminal Protein Localized to the Nucleus and Presynaptic Nerve Terminal. *J. Neurosci.* 1988;8(8):2804-15.
84. Scherzer CR, Grass JA, Liao Z, Pepivani I, Zheng B, Eklund AC, *et al.* GATA transcription factors directly regulate the Parkinson ' s disease-linked gene alpha-synuclein. *Proc. Natl. Acad. Sci. U.S.A.* 2008;105(31):10907-12.
85. Chiba-Falek O, Nussbaum RL. Effect of allelic variation at the NACP-Rep1 repeat upstream of the alpha-synuclein gene (SNCA) on transcription in a cell culture luciferase reporter system. *Hum. Mol. Genet.* 2001;10(26):3101-9.
86. Markopoulou K, Biernacka JM, Armasu SM, Anderson KJ, Ahlskog JE, Chase BA, *et al.* Does  $\alpha$ -synuclein have a dual and opposing effect in preclinical vs. clinical Parkinson's disease? *Parkinsonism. Relat. Disord.* 2014;20(6):584-9.
87. Cronin KD, Ge D, Manninger P, Linnertz C, Rossoshek A, Orrison BM, *et al.* Expansion of the Parkinson disease-associated SNCA-Rep1 allele upregulates human  $\alpha$ -synuclein in transgenic mouse brain. *Hum. Mol. Genet.* 2009;18(17):3274-85.
88. Sotiriou S, Gibney G, Baxevanis AD, Nussbaum RL. A single nucleotide polymorphism in the 3'UTR of the SNCA gene encoding alpha-synuclein is a new potential susceptibility locus for Parkinson disease. *Neurosci. Lett.* 2009;461(2):196-201.
89. Yu L, Xu P, He X, Hu F, Lin Z, Zhu M, *et al.* SNP rs7684318 of the alpha-synuclein gene is associated with Parkinson's disease in the Han Chinese population. *Brain. Res.* 2010;1346:262-5.
90. Tan YY, Wu L, Zhao ZB, Wang Y, Xiao Q, Liu J, *et al.* Methylation of  $\alpha$ -synuclein and leucine-rich repeat kinase 2 in leukocyte DNA of Parkinson's disease patients. *Parkinsonism. Relat. Disord.* 2014;20(3):308-13.
91. Matsumoto L, Takuma H, Tamaoka A, Kurisaki H, Date H, Tsuji S, *et al.* CpG demethylation enhances alpha-synuclein expression and affects the pathogenesis of Parkinson's disease. *PLoS One.* 2010;5(11):e15522.
92. Jowaed A, Schmitt I, Kaut O, Wüllner U. Methylation regulates alpha-synuclein expression and is decreased in Parkinson's disease patients' brains. *J. Neurosci.* 2010;30(18):6355-9.
93. Ai SX, Xu Q, Hu YC, Song CY, Guo JF, Shen L, *et al.* Hypomethylation of SNCA in blood of



- patients with sporadic Parkinson's disease. *J. Neurol. Sci.* 2014;337(1-2):123-8.
94. Desplats P, Spencer B, Coffee E, Patel P, Michael S, Patrick C, *et al.*  $\alpha$ -Synuclein Sequesters Dnmt1 from the Nucleus. *J. Biol. Chem.* 2011; 286(11):9031-7.
  95. Renella R, Schlehe JS, Selkoe DJ, Williams DA, LaVoie MJ. Genetic deletion of the GATA1-regulated protein  $\alpha$ -synuclein reduces oxidative stress and nitric oxide synthase levels in mature erythrocytes. *Am. J. Hematol.* 2014;89(10):974-7.
  96. Junn E, Lee KW, Jeong BS, Chan TW, Im JY, Mouradian MM. Repression of alpha-synuclein expression and toxicity by microRNA-7. *Proc. Natl. Acad. Sci. U. S. A.* 2009;106(31):13052-7.
  97. Schmitt I, Wüllner U, van Rooyen JP, Khazneh H, Becker J, Volk A, *et al.* Variants in the 3'UTR of SNCA do not affect miRNA-433 binding and alpha-synuclein expression. *Eur. J. Hum. Genet.* 2012;20(12):1265-9.
  98. Doxakis E. Post-transcriptional regulation of a-synuclein expression by mir-7 and mir-153. *J. Biol. Chem.* 2010;285(17):12726-34.
  99. Stefanis L.  $\alpha$ -Synuclein in Parkinson's Disease. *Cold. Spring. Harb. Perspect. Med.* 2012;2(2):a009399.
  100. Wood S, Wypych J, Steavenson S, Louis JC, Citron M, Biere AL. alpha -Synuclein Fibrillogenesis
  101. Narhi, L. *et al.* Both familial Parkinson's disease mutations accelerate alpha-synuclein aggregation. *J. Biol. Chem.* 274, 9843–9846 (1999).
  102. Hashimoto, M., Takeda, a, Hsu, L. J., Takenouchi, T. & Masliah, E. Role of cytochrome c as a stimulator of alpha-synuclein aggregation in Lewy body disease. *J. Biol. Chem.* 274, 28849–28852 (1999).
  103. Nishioka, K. *et al.* Clinical heterogeneity of alpha-synuclein gene duplication in Parkinson's disease. *Ann. Neurol.* 59, 298–309 (2006).
  104. Kenya Nishioka, O. a. R. and N. H. Gene Duplication. (2011). doi:10.5772/888
  105. Murphy, D. D., Rueter, S. M., Trojanowski, J. Q. & Lee, V. M. Synucleins Are Developmentally Expressed , and  $\alpha$ -Synuclein Regulates the Size of the Presynaptic Vesicular Pool in Primary Hippocampal Neurons. 20, 3214–3220 (2000).
  106. Abeliovich, A. *et al.* Mice Lacking  $\alpha$ -Synuclein Display Functional Deficits in the Nigrostriatal Dopamine System. 25, 239–252 (2000).
  107. Bonini, N. M. & Giasson, B. I. Snaring the function of ??-synuclein. *Cell* 123, 359–361 (2005).
  108. Burré, J., Sharma, M. & Südhof, T. C. upon membrane binding to promote SNARE. (2014). doi:10.1073/pnas.1416598111
  109. Fusion, D. V. HHS Public Access. 54, 1831–1840 (2015).
  110. Davidson, W. S. Stabilization of alpha -Synuclein Secondary Structure upon Binding to Synthetic Membranes. *J. Biol. Chem.* 273, 9443–9449 (1998).
  111. Jensen, P. H., Nielsen, M. S., Jakes, R., Dotti, C. G. & Goedert, M. Binding of  $\alpha$ -Synuclein to Brain Vesicles Is Abolished by Familial Parkinson's Disease Mutation. *J. Biol. Chem.* 273, 26292–26294 (1998).
  112. Tanaka, Y. *et al.* Inducible expression of mutant  $\alpha$  -synuclein decreases proteasome activity and increases sensitivity to mitochondria-dependent apoptosis. 10, 919–926 (2001).
  113. Cook, C. *et al.* Disruption of Protein Quality Control in Parkinson ' s Disease. 1–18 (2013).
  114. Mak, S. K., McCormack, A. L., Manning-Bog, A. B., Cuervo, A. M. & Di Monte, D. a. Lysosomal degradation of  $\alpha$ -synuclein in vivo. *J. Biol. Chem.* 285, 13621–13629 (2010).

115. Manning-Bog, A. B. et al. The herbicide paraquat causes up-regulation and aggregation of ??-synuclein in mice: Paraquat and ??-synuclein. *J. Biol. Chem.* 277, 1641–1644 (2002).
116. Ciron, C. et al. PGC-1 $\alpha$  activity in nigral dopamine neurons determines vulnerability to  $\alpha$ -synuclein. *Acta Neuropathol. Commun.* 3, 1–20 (2015).
117. Nichols, W. C. et al. Genetic screening for a single common LRRK2 mutation in familial Parkinson's disease. *Lancet* 365, 410–412 (2005).
118. Di Fonzo, A. et al. A frequent LRRK2 gene mutation associated with autosomal dominant Parkinson's disease. *Lancet* 365, 412–415 (2005).
119. Sierra, M. et al. High frequency and reduced penetrance of LRRK2 G2019S mutation among Parkinson's disease patients in Cantabria (Spain). *Mov. Disord.* 26, 2343–2346 (2011).
120. Gaig, C. et al. G2019S LRRK2 mutation causing Parkinson's disease without Lewy bodies. *J. Neurol. Neurosurg. Psychiatry* 78, 626–628 (2007).
121. Elliott, D. a, Kim, W. S., Gorissen, S., Halliday, G. M. & Kwok, J. B. J. Leucine-rich repeat kinase 2 and alternative splicing in Parkinson's disease. *Mov. Disord.* 27, 1004–11 (2012).
122. Marras, C. et al. Phenotype in parkinsonian and nonparkinsonian LRRK2 G2019S mutation carriers. *Neurology* 77, 325–333 (2011).
123. West, A. B. et al. Parkinson's disease-associated mutations in leucine-rich repeat kinase 2 augment kinase activity. *Proc. Natl. Acad. Sci. U. S. A.* 102, 16842–16847 (2005).
124. Cho, H. J. et al. MicroRNA-205 regulates the expression of parkinson's disease-related leucine-rich repeat kinase 2 protein. *Hum. Mol. Genet.* 22, 608–620 (2013).
125. Kachergus, J. et al. Identification of a novel LRRK2 mutation linked to autosomal dominant parkinsonism: evidence of a common founder across European populations. *Am. J. Hum. Genet.* 76, 672–680 (2005).
126. Mata, I. F., Wedemeyer, W. J., Farrer, M. J., Taylor, J. P. & Gallo, K. a. LRRK2 in Parkinson's disease: protein domains and functional insights. *Trends Neurosci.* 29, 286–293 (2006).
127. Cirnaru, M. D. et al. LRRK2 kinase activity regulates synaptic vesicle trafficking and neurotransmitter release through modulation of LRRK2 macro-molecular complex. *Front. Mol. Neurosci.* 7, 49 (2014).
128. Piccoli, G. et al. Leucine-Rich Repeat Kinase 2 Binds to Neuronal Vesicles through Protein Interactions Mediated by Its C-Terminal WD40 Domain. *Mol. Cell. Biol.* 34, 2147–2161 (2014).
129. Hatano, T. et al. Leucine-rich repeat kinase 2 associates with lipid rafts. *Hum. Mol. Genet.* 16, 678–690 (2007).
130. Li, J., Tan, L. & Yu, J. The role of the LRRK2 gene in Parkinsonism. 1–17 (2014).
131. MacLeod, D. et al. The Familial Parkinsonism Gene LRRK2 Regulates Neurite Process Morphology. *Neuron* 52, 587–593 (2006).
132. Manuscript, A. Role of autophagy in G2019S-LRRK2-associated neurite shortening in differentiated SH-SY5Y cells. *Changes* 29, 997–1003 (2012).
133. Bravo-San Pedro, J. M. et al. The MAPK1/3 pathway is essential for the deregulation of autophagy observed in G2019S LRRK2 mutant fibroblasts. *Autophagy* 8, 1537–1539 (2012).
134. Tong, Y. et al. Loss of leucine-rich repeat kinase 2 causes impairment of protein degradation pathways, accumulation of alpha-synuclein, and apoptotic cell death in aged mice. *Proc. Natl. Acad. Sci. U. S. A.* 107, 9879–9884 (2010).
135. Hsu, C. H., Chan, D. & Wolozin, B. LRRK2 and the stress response: Interaction with MKKs and JNK-interacting proteins. *Neurodegener. Dis.* 7, 68–75 (2010).

136. Gandhi, P. N., Wang, X., Zhu, X., Chen, S. G. & Wilson-Delfosse, A. L. The Roc domain of leucine-rich repeat kinase 2 is sufficient for interaction with microtubules. *J. Neurosci. Res.* 86, 1711–1720 (2008).
137. Liou, A. K. F., Leak, R. K., Li, L. & Zigmond, M. J. Wild-type LRRK2 but not its mutant attenuates stress-induced cell death via ERK pathway. *Neurobiol. Dis.* 32, 116–124 (2008).
138. Yuan, Y. et al. Dysregulated LRRK2 signaling in response to endoplasmic reticulum stress leads to dopaminergic neuron degeneration in *C. elegans*. *PLoS One* 6, (2011).
139. Imai, Y. et al. Phosphorylation of 4E-BP by LRRK2 affects the maintenance of dopaminergic neurons in *Drosophila*. *EMBO J.* 27, 2432–2443 (2008).
140. Russo, I., Bubacco, L. & Greggio, E. LRRK2 and neuroinflammation: partners in crime in Parkinson's disease? *J. Neuroinflammation* 11, 52 (2014).
141. Marker, D. F. et al. LRRK2 kinase inhibition prevents pathological microglial phagocytosis in response to HIV-1 Tat protein. *J. Neuroinflammation* 9, 261 (2012).
142. Rivas, M. a et al. Variants Associated With Inflammatory Bowel Disease. *Nat. Genet.* 43, 1066–1073 (2012).
143. Lu, C.-S., Wu Chou, Y.-H., Kuo, P.-C., Chang, H.-C. & Weng, Y.-H. The parkinsonian phenotype of spinocerebellar ataxia type 2. *Arch. Neurol.* 61, 35–8 (2004).
144. Magaña, J. J., Velázquez-Pérez, L. & Cisneros, B. Spinocerebellar ataxia type 2: Clinical presentation, molecular mechanisms, and therapeutic perspectives. *Molecular Neurobiology* 47, 90–104 (2013).
145. Elden, A. A. C. et al. Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. *Nature* 466, 1069–1075 (2010).
146. Wang, L. et al. Large-scale assessment of polyglutamine repeat expansions in Parkinson disease. 1283–1292 (2015).
147. Gehman, L. T. et al. The splicing regulator Rbfox1 (A2BP1) controls neuronal excitation in the mammalian brain. *Nat. Genet.* 43, 706–711 (2011).
148. Huynh, D. P., Del Bigio, M. R., Ho, D. H. & Pulst, S. M. Expression of ataxin-2 in brains from normal individuals and patients with Alzheimer's disease and spinocerebellar ataxia 2. *Ann. Neurol.* 45, 232–241 (1999).
149. Laffita-Mesa, J. M. et al. Epigenetics DNA methylation in the core ataxin-2 gene promoter: novel physiological and pathological implications. *Hum. Genet.* 131, 625–38 (2012).
150. Jiménez-López, D. & Guzmán, P. Insights into the evolution and domain structure of Ataxin-2 proteins across eukaryotes. *BMC Res. Notes* 7, 453 (2014).
151. He, W. & Parker, R. Functions of Lsm proteins in mRNA degradation and splicing. *Curr. Opin. Cell Biol.* 12, 346–350 (2000).
152. van de Loo, S., Eich, F., Nonis, D., Auburger, G. & Nowock, J. Ataxin-2 associates with rough endoplasmic reticulum. *Exp. Neurol.* 215, 110–118 (2009).
153. Zhang, F. et al. Association analysis of a functional variant in ATXN2 with schizophrenia. *Neurosci. Lett.* 562, 24–7 (2014).
154. Welzel, F. et al. FOX-2 dependent splicing of ataxin-2 transcript is affected by ataxin-1 overexpression. *PLoS ONE* 7, (2012).
155. Eliseeva, I. a, Lyabin, D. N. & Ovchinnikov, L. P. Poly(A)-binding proteins: structure, domain organization, and activity regulation. *Biochem. Biokhimiia* 78, 1377–91 (2013).
156. Yokoshi, M. et al. Direct Binding of Ataxin-2 to Distinct Elements in 3' UTRs Promotes mRNA

- Stability and Protein Expression. *Mol. Cell* 55, 186–198 (2014).
157. Fittschen, M. et al. Genetic ablation of ataxin-2 increases several global translation factors in their transcript abundance but decreases translation rate. *Neurogenetics* 16, 181–192 (2015).
  158. Wang, Y., Arribas-Layton, M., Chen, Y., Lykke-Andersen, J. & Sen, G. L. DDX6 Orchestrates Mammalian Progenitor Function through the mRNA Degradation and Translation Pathways. *Mol. Cell* 60, 118–130 (2015).
  159. Satterfield, T. F. & Pallanck, L. J. Ataxin-2 and its *Drosophila* homolog, ATX2, physically assemble with polyribosomes. *Hum. Mol. Genet.* 15, 2523–32 (2006).
  160. Mangus, D. a, Amrani, N. & Jacobson, a. Pbp1p, a factor interacting with *Saccharomyces cerevisiae* poly(A)-binding protein, regulates polyadenylation. *Mol. Cell. Biol.* 18, 7383–7396 (1998).
  161. Kozlov, G., Safaee, N., Rosenauer, A. & Gehring, K. Structural basis of binding of P-body-associated proteins GW182 and ataxin-2 by the Mille domain of poly(A)-binding protein. *J. Biol. Chem.* 285, 13599–13606 (2010).
  162. Ralser, M. et al. Ataxin-2 and huntingtin interact with endophilin-A complexes to function in plastin-associated pathways. *Hum. Mol. Genet.* 14, 2893–2909 (2005).
  163. Lastres-Becker, I. et al. Insulin receptor and lipid metabolism pathology in ataxin-2 knock-out mice. *Hum. Mol. Genet.* 17, 1465–1481 (2008).
  164. Velázquez-Pérez, L. et al. A comprehensive review of spinocerebellar ataxia type 2 in Cuba. *Cerebellum* 10, 184–98 (2011).
  165. Scuderi, S., La Cognata, V., Drago, F., Cavallaro, S. & D’Agata, V. Alternative splicing generates different parkin protein isoforms: evidences in human, rat, and mouse brain. *Biomed Res. Int.* 2014, 690796 (2014).
  166. Periquet, M. et al. Origin of the mutations in the parkin gene in Europe: exon rearrangements are independent recurrent events, whereas point mutations may result from Founder effects. *Am. J. Hum. Genet.* 68, 617–626 (2001).
  167. Seirafi, M., Kozlov, G. & Gehring, K. Parkin Structure and Function. *FEBS J.* 282, 1–13 (2015).
  168. Wauer, T. & Komander, D. Structure of the human Parkin ligase domain in an autoinhibited state. *EMBO J.* 32, 2099–112 (2013).
  169. McLelland, G. L., Soubannier, V., Chen, C. X., McBride, H. M. & Fon, E. a. Parkin and PINK1 function in a vesicular trafficking pathway regulating mitochondrial quality control. *EMBO J.* 33, 282–295 (2014).
  170. Shin, J.-H. et al. PARIS (ZNF746) repression of PGC-1 $\alpha$  contributes to neurodegeneration in Parkinson’s disease. *Cell* 144, 689–702 (2011).
  171. Valente, E. M. et al. Hereditary early-onset Parkinson’s disease caused by mutations in PINK1. *Science* 304, 1158–1160 (2004).
  172. Matsuda, S., Kitagishi, Y. & Kobayashi, M. Function and characteristics of PINK1 in mitochondria. *Oxid. Med. Cell. Longev.* 2013, (2013).
  173. Dzamko, N., Zhou, J., Huang, Y. & Halliday, G. M. Parkinson’s disease-implicated kinases in the brain; insights into disease pathogenesis. *Front. Mol. Neurosci.* 7, 57 (2014).
  174. Ashrafi, G., Schlehe, J. S., LaVoie, M. J. & Schwarz, T. L. Mitophagy of damaged mitochondria occurs locally in distal neuronal axons and requires PINK1 and Parkin. *J. Cell Biol.* 206, 655–670 (2014).
  175. Morais, V. a. et al. Parkinson’s disease mutations in PINK1 result in decreased Complex I activity and deficient synaptic function. *EMBO Mol. Med.* 1, 99–111 (2009).

176. Song, S. et al. Characterization of PINK1 (PTEN-induced putative kinase 1) mutations associated with parkinson disease in mammalian cells and drosophila. *J. Biol. Chem.* 288, 5660–5672 (2013).
177. Fedorowicz, M. a. et al. Cytosolic cleaved PINK1 represses Parkin translocation to mitochondria and mitophagy. *EMBO Rep.* 15, 86–93 (2014).
178. Pickrell, A. M. & Youle, R. J. The Roles of PINK1, Parkin, and Mitochondrial Fidelity in Parkinson's Disease. *Neuron* 85, 257–273 (2015).
179. Bonifati, V. et al. DJ-1( PARK7), a novel gene for autosomal recessive, early onset parkinsonism. *Neurol. Sci.* 24, 159–160 (2003).
180. Giaime, E., Yamaguchi, H., Gautier, C. a., Kitada, T. & Shen, J. Loss of DJ-1 does not affect mitochondrial respiration but increases ROS production and mitochondrial permeability transition pore opening. *PLoS One* 7, (2012).
181. Mullett, S. J. & Hinkle, D. a. DJ-1 deficiency in astrocytes selectively enhances mitochondrial Complex i inhibitor-induced neurotoxicity. *J. Neurochem.* 117, 375–387 (2011).
182. Zhou, W. & Freed, C. R. DJ-1 up-regulates glutathione synthesis during oxidative stress and inhibits A53T ??-synuclein toxicity. *J. Biol. Chem.* 280, 43150–43158 (2005).
183. McCoy, M. K. & Cookson, M. R. DJ-1 regulation of mitochondrial function and autophagy through oxidative stress. *Autophagy* 7, 531–532 (2011).
184. Choi, D.-H. et al. DJ-1 cleavage by matrix metalloproteinase 3 mediates oxidative stress-induced dopaminergic cell death. *Antioxid. Redox Signal.* 14, 2137–2150 (2011).
185. Giroto, S. et al. DJ-1 Is a copper chaperone acting on SOD1 activation. *J. Biol. Chem.* 289, 10887–10899 (2014).
186. Kamp, F. et al. Inhibition of mitochondrial fusion by  $\alpha$ -synuclein is rescued by PINK1, Parkin and DJ-1. *EMBO J.* 29, 3571–3589 (2010).
187. Mccoy, M. K. & Cookson, M. R. DJ-1 regulation of mitochondrial function and autophagy through oxidative stress. *Autophagy* 1–2 (2011). doi:10.1093/hmg/ddq430.he
188. Ishikawa, S. et al. Human DJ-1-specific transcriptional activation of tyrosine hydroxylase gene. *J. Biol. Chem.* 285, 39718–39731 (2010).
189. Kim, K. S. et al. DJ-1 associates with lipid rafts by palmitoylation and regulates lipid rafts-dependent endocytosis in astrocytes. *Hum. Mol. Genet.* 22, 4805–4817 (2013).
190. Ren, H., Fu, K., Mu, C., Zhen, X. & Wang, G. L166P mutant DJ-1 promotes cell death by dissociating Bax from mitochondrial Bcl-XL. *Mol. Neurodegener.* 7, 40 (2012).
191. Vilariño-Güell, C. et al. VPS35 mutations in parkinson disease. *Am. J. Hum. Genet.* 89, 162–167 (2011).
192. Zimprich, A. et al. A mutation in VPS35, encoding a subunit of the retromer complex, causes late-onset parkinson disease. *Am. J. Hum. Genet.* 89, 168–175 (2011).
193. Manuscript, A. & Manuscript, A. The retromer subunit Vps26 has an arrestin fold and binds Vps35 through its C-terminal domain. *Structure* 13, 540–548 (2006).
194. Wen, L. et al. VPS35 haploinsufficiency increases Alzheimer's disease neuropathology. *J. Cell Biol.* 195, 765–779 (2011).
195. Korolchuk, V. I. et al. Drosophila Vps35 function is necessary for normal endocytic trafficking and actin cytoskeleton organisation. *J. Cell Sci.* 120, 4367–4376 (2007).
196. Braschi, E. et al. Vps35 mediates vesicle transport between the mitochondria and peroxisomes. *Curr. Biol.* 20, 1310–1315 (2010).

197. Bi, F., Li, F., Huang, C. & Zhou, H. Pathogenic mutation in VPS35 impairs its protection against MPP+ cytotoxicity. *Int. J. Biol. Sci.* 9, 149–155 (2013).
198. Malik, B. R., Godena, V. K. & Whitworth, A. J. VPS35 pathogenic mutations confer no dominant toxicity but partial loss of function in *Drosophila* and genetically interact with parkin. *Hum. Mol. Genet.* ddv322 (2015). doi:10.1093/hmg/ddv322
199. Miura, E. et al. VPS35 dysfunction impairs lysosomal degradation of  $\alpha$ -synuclein and exacerbates neurotoxicity in a *Drosophila* model of Parkinson's disease. *Neurobiol. Dis.* 71, 1–13 (2014).
200. Chartier-Harlin, M. C. et al. Translation initiator EIF4G1 mutations in familial parkinson disease. *Am. J. Hum. Genet.* 89, 398–406 (2011).
201. Schulte, E. C. et al. Variants in eukaryotic translation initiation factor 4G1 in sporadic Parkinson's disease. *Neurogenetics* 13, 281–285 (2012).
202. Nichols, N. et al. EIF4G1 mutations do not cause Parkinson's disease. *Neurobiol. Aging* 36, 2444.e1–2444.e4 (2015).
203. Deng, H., Wu, Y. & Jankovic, J. The EIF4G1 gene and Parkinson's disease. *Acta Neurol. Scand.* 1, n/a–n/a (2015).
204. Ryu, I., Park, J. H., An, S., Kwon, O. S. & Jang, S. K. eIF4G1 Facilitates the MicroRNA-Mediated Gene Silencing. *PLoS One* 8, (2013).
205. Dhungel, N. et al. Parkinson's Disease Genes VPS35 and EIF4G1 Interact Genetically and Converge on  $\alpha$ -Synuclein. *Neuron* 85, 76–87 (2015).
206. Erro, R. VPS35 and EIF4G1 interactions and novel candidate genes for PD: From genes to pathways and back. *Mov. Disord.* 30, 499–499 (2015).
207. Huttenlocher, J. et al. EIF4G1 is neither a strong nor a common risk factor for Parkinson's disease: evidence from large European cohorts: Table1. *J. Med. Genet.* 52, 37–41 (2015).
208. Simon-Sanchez, J., Schulte, C. & Bras, J. Genome-wide association study reveals genetic risk underlying Parkinson's disease. ... *Genet.* 41, 1308–1312 (2009).
209. Pankratz, N. Genomewide association study for susceptibility genes contributing to familial Parkinson disease. *124*, 593–605 (2010).
210. Edwards, T. L. et al. Genome-wide association study confirms SNPs in SNCA and the MAPT region as common risk factors for Parkinson disease. *74*, 97–109 (2010).
211. Tan, E. K. et al. Analysis of GWAS-linked loci in Parkinson disease reaffirms PARK16 as a susceptibility locus. *Neurology* 75, 508–512 (2010).
212. Spencer, C. C. a et al. Dissection of the genetics of Parkinson's disease identifies an additional association 5' of SNCA and multiple associated haplotypes at 17q21. *Hum. Mol. Genet.* 20, 345–353 (2011).
213. Evangelou, E., Maraganore, D. M. & Ioannidis, J. P. a. Meta-analysis in genome-wide association datasets: Strategies and application in Parkinson disease. *PLoS One* 2, (2007).
214. D, P., Doheny, K. F., Factor, S. a & H, T. Meta-analysis of Parkinson diseases: Identification of a novel locus, RIT2. *71*, 370–384 (2013).
215. Rhodes, S. L., Sinsheimer, J. S., Bordelon, Y., Bronstein, J. M. & Ritz, B. Replication of GWAS Associations for GAK and MAPT in Parkinson's Disease. *Ann. Hum. Genet.* 75, 195–200 (2011).
216. Hamza, T. H. et al. NIH Public Access. *42*, 781–785 (2011).
217. Nalls, M. a. et al. Genetic comorbidities in parkinson's disease. *Hum. Mol. Genet.* 23, 831–841 (2014).

218. Liu, X. et al. Genome-Wide association study identifies candidate genes for Parkinson's disease in an Ashkenazi Jewish population. *BMC Med. Genet.* 12, 104 (2011).
219. Jowaed, A., Schmitt, I., Kaut, O. & Wüllner, U. Methylation regulates alpha-synuclein expression and is decreased in Parkinson's disease patients' brains. *J. Neurosci.* 30, 6355–9 (2010).
220. Goers, J. et al. Nuclear localization of alpha-synuclein and its interaction with histones. *Biochemistry* 42, 8465–8471 (2003).
221. Outeiro, T. F. et al. Models of Parkinson's Disease. 1968, (2007).
222. Siddiqui, A. et al. Selective binding of nuclear alpha-synuclein to the PGC1alpha promoter under conditions of oxidative stress may contribute to losses in mitochondrial function: Implications for Parkinson's disease. *Free Radic. Biol. Med.* 53, 993–1003 (2012).
223. Gehrke, S., Imai, Y., Sokol, N. & Lu, B. Pathogenic LRRK2 negatively regulates microRNA-mediated translational repression. *Nature* 466, 637–41 (2010).
224. Verstraeten, A., Theuns, J. & Van Broeckhoven, C. Progress in unraveling the genetic etiology of Parkinson disease in a genomic era. *Trends Genet.* 31, 140–149 (2015).
225. Grünblatt, E. et al. Gene expression profiling of parkinsonian substantia nigra pars compacta; alterations in ubiquitin-proteasome, heat shock protein, iron and oxidative stress regulated proteins, cell adhesion/cellular matrix and vesicle trafficking genes. *J. Neural Transm.* 111, 1543–1573 (2004).
226. Duke, D. C. et al. Transcriptome analysis reveals link between proteasomal and mitochondrial pathways in Parkinson's disease. *Neurogenetics* 7, 139–148 (2006).
227. Simunovic, F. et al. Gene expression profiling of substantia nigra dopamine neurons: further insights into Parkinson's disease pathology. *Brain* 132, 1795–1809 (2009).
228. Torrent, R. et al. Using iPSC Cells toward the Understanding of Parkinson's Disease. *J. Clin. Med.* 4, 548–566 (2015).
229. Woodard, C. M. et al. iPSC-Derived Dopamine Neurons Reveal Differences between Monozygotic Twins Discordant for Parkinson's Disease. *Cell Rep.* 9, 1173–1182 (2014).
230. Auburger, G. et al. Primary skin fibroblasts as a model of Parkinson's disease. *Mol. Neurobiol.* 46, 20–27 (2012).
231. Hoepken, H.-H. et al. Parkinson patient fibroblasts show increased alpha-synuclein expression. *Exp. Neurol.* 212, 307–13 (2008).
232. Barroso, N. et al. Respiratory chain enzyme activities in lymphocytes from untreated patients with Parkinson disease. *Clin Chem* 39, 667–669 (1993).
233. Molina, J. a et al. Respiratory chain enzyme activities in isolated mitochondria of lymphocytes from patients with Alzheimer's disease. *Neurology* 48, 636–638 (1997).
234. Year, I. C. Mitochondrial complex - 1 in Parkinson's disease. 3–4 (2015).
235. Prigione, A. et al. Oxidative stress in peripheral blood mononuclear cells from patients with Parkinson's disease: Negative correlation with levodopa dosage. *Neurobiol. Dis.* 23, 36–43 (2006).
236. Brochard, V. et al. Infiltration of CD4+ lymphocytes into the brain contributes to neurodegeneration in a mouse model of Parkinson disease. *J. Clin. Invest.* 119, 182–192 (2009).
237. Grozdanov, V. et al. Inflammatory dysregulation of blood monocytes in Parkinson's disease patients. *Acta Neuropathol.* 128, 651–663 (2014).

238. Bessler, H., Djaldetti, R., Salman, H., Bergman, M. & Djaldetti, M. IL-1 beta, IL-2, IL-6 and TNF-alpha production by peripheral blood mononuclear cells from patients with Parkinson's disease. *Biomed. Pharmacother.* 53, 141–145 (1999).
239. Ricci, a et al. Dopamine D3 receptor in peripheral mononuclear cells of essential hypertensives. *Hypertension* 30, 1566–1571 (1997).
240. Scherzer, C. R. et al. Molecular markers of early Parkinson's disease based on gene expression in blood. *Proceedings of the National Academy of Sciences of the United States of America* 104, 955–960 (2007).
241. Mutez, E. et al. Transcriptional profile of Parkinson blood mononuclear cells with LRRK2 mutation. *Neurobiol. Aging* 32, 1839–1848 (2011).
242. Botta-Orfila, T. et al. Brain transcriptomic profiling in idiopathic and LRRK2-associated Parkinson's disease. *Brain Res.* 1466, 152–157 (2012).
243. Diao, H., Li, X. & Hu, S. The identification of dysfunctional crosstalk of pathways in Parkinson disease. *Gene* 515, 159–162 (2013).
244. Infante, J. et al. Identification of candidate genes for Parkinson's disease through blood transcriptome analysis in LRRK2-G2019S carriers, idiopathic cases, and controls. *Neurobiol. Aging* 36, 1105–1109 (2015).
245. Sambrook, J. Adenovirus amazes at Cold Spring Harbor. *Nature* 268, 101–104 (1977).
246. Keren, H., Lev-Maor, G. & Ast, G. Alternative splicing and evolution: diversification, exon definition and function. *Nat. Rev. Genet.* 11, 345–55 (2010).
247. Patel, S. B. & Bellini, M. The assembly of a spliceosomal small nuclear ribonucleoprotein particle. *Nucleic Acids Res.* 36, 6482–6493 (2008).
248. Patel, A. a & Steitz, J. a. Splicing double: insights from the second spliceosome. *Nat. Rev. Mol. Cell Biol.* 4, 960–970 (2003).
249. Will, C. L. & Lührmann, R. Spliceosome structure and function. *Cold Spring Harb. Perspect. Biol.* 3, 1–23 (2011).
250. Ramchatesingh, J., Zahler, a M., Neugebauer, K. M., Roth, M. B. & Cooper, T. a. A subset of SR proteins activates splicing of the cardiac troponin T alternative exon by direct interactions with an exonic enhancer. *Mol. Cell. Biol.* 15, 4898–4907 (1995).
251. Jamison, S. F. et al. U1 snRNP-ASF/SF2 interaction and 5' splice site recognition: characterization of required elements. *Nucleic Acids Res.* 23, 3260–3267 (1995).
252. Roscigno, R. F. & Garcia-Blanco, M. a. SR proteins escort the U4/U6.U5 tri-snRNP to the spliceosome. *RNA (New York, N.Y.)* 1, 692–706 (1995).
253. Cao, W. & Garcia-Blanco, M. a. A serine/arginine-rich domain in the human U1 70k protein is necessary and sufficient for ASF/SF2 binding. *J. Biol. Chem.* 273, 20629–20635 (1998).
254. Chan, C. B. & Ye, K. Serine-arginine protein kinases: New players in neurodegenerative diseases? *Rev. Neurosci.* 24, 401–413 (2013).
255. Blaustein, M., Pelisch, F. & Srebrow, A. Signals, pathways and splicing regulation. *Int. J. Biochem. Cell Biol.* 39, 2031–2048 (2007).
256. Qian, W. et al. Regulation of the alternative splicing of tau exon 10 by SC35 and Dyrk1A. *Nucleic Acids Res.* 39, 6161–6171 (2011).
257. White, E. S. et al. Control of fibroblast fibronectin expression and alternative splicing via the PI3K/Akt/mTOR pathway. *Exp. Cell Res.* 316, 2644–53 (2010).
258. Dudek, H. et al. Regulation of neuronal survival by the serine-threonine protein kinase Akt.



- Science 275, 661–5 (1997).
259. Shultz, J. C. et al. Alternative splicing of caspase 9 is modulated by the phosphoinositide 3-kinase/Akt pathway via phosphorylation of SRp30a. *Cancer Res.* 70, 9185–9196 (2010).
  260. Shultz, J. C. et al. Control of fibroblast fibronectin expression and alternative splicing via the PI3K/Akt/mTOR pathway. *Cancer Res.* 70, 9185–9196 (2010).
  261. Sacco-Bubulya, P. & Spector, D. L. Disassembly of interchromatin granule clusters alters the coordination of transcription and pre-mRNA splicing. *J. Cell Biol.* 156, 425–436 (2002).
  262. Lamond, A. I. & Spector, D. L. Nuclear speckles: a model for nuclear organelles. *Nat. Rev. Mol. Cell Biol.* 4, 605–612 (2003).
  263. Cáceres, J. F., Misteli, T., Sreaton, G. R., Spector, D. L. & Krainer, A. R. Role of the modular domains of SR proteins in subnuclear localization and alternative splicing specificity. *J. Cell Biol.* 138, 225–238 (1997).
  264. Clelland, a. K., Bales, a. B. E. & Sleeman, J. E. Changes in intranuclear mobility of mature snRNPs provide a mechanism for splicing defects in spinal muscular atrophy. *J. Cell Sci.* 125, 2626–2637 (2012).
  265. McManus, C. J. & Graveley, B. R. RNA structure and the mechanisms of alternative splicing. *Curr. Opin. Genet. Dev.* 21, 373–379 (2011).
  266. Varani, L. et al. Structure of tau exon 10 splicing regulatory element RNA and destabilization by mutations of frontotemporal dementia and parkinsonism linked to chromosome 17. *Proc. Natl. Acad. Sci. U. S. A.* 96, 8229–8234 (1999).
  267. Buratti, E. & Baralle, F. E. Influence of RNA Secondary Structure on the Pre-mRNA Splicing Process MINIREVIEW Influence of RNA Secondary Structure on the Pre-mRNA Splicing Process. *Mol. Cell. Biol.* 24, 10505–10514 (2004).
  268. Kalsotra, A., Wang, K., Li, P. & Cooper, T. a. MicroRNAs coordinate an alternative splicing network during mouse postnatal heart development Email alerting service MicroRNAs coordinate an alternative splicing network during mouse postnatal heart development. *Genes Dev.* 24, 653–658 (2010).
  269. Boutz, P. L., Chawla, G., Stoilov, P. & Black, D. L. MicroRNAs regulate the expression of the alternative splicing factor nPTB during muscle development. *Genes Dev.* 21, 71–84 (2007).
  270. Tripathi, V. et al. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol. Cell* 39, 925–938 (2010).
  271. Cáceres, J. F. & Kornblihtt, A. R. Alternative splicing: Multiple control mechanisms and involvement in human disease. *Trends in Genetics* 18, 186–193 (2002).
  272. Seng, C. O., Magee, C., Young, P. J., Lorson, C. L. & Allen, J. P. The SMN structure reveals its crucial role in snRNP assembly. *Hum. Mol. Genet.* 24, 2138–2146 (2015).
  273. Elliott, D. A., Kim, W. S., Gorissen, S., Halliday, G. M. & Kwok, J. B. J. Leucine-rich repeat kinase 2 and alternative splicing in Parkinson’s disease. *Mov. Disord.* 27, 1004–1011 (2012).
  274. Di Fonzo, A. et al. Comprehensive analysis of the LRRK2 gene in sixty families with Parkinson’s disease. *Eur. J. Hum. Genet.* 14, 322–31 (2006).
  275. Kalivendi, S. V, Yedlapudi, D., Hillard, C. J., Kalyanaraman, B. & Road, U. Oxidants induce alternative splicing of a-synuclein: implications for Parkinson’s disease. 48, 377–383 (2015).
  276. Bungereth, M. et al. Differential aggregation properties of alpha-synuclein isoforms. *Neurobiol. Aging* 35, 1913–1919 (2014).
  277. Alieva, A. K. et al. Involvement of endocytosis and alternative splicing in the formation of the pathological process in the early stages of parkinson’s disease. *Biomed Res. Int.* 2014, (2014).

278. Shehadeh, L. a. et al. SRRM2, a potential blood biomarker revealing high alternative splicing in Parkinson's disease. *PLoS One* 5, 1–8 (2010).
279. Elstner, M. et al. Expression analysis of dopaminergic neurons in Parkinson's disease and aging links transcriptional dysregulation of energy metabolism to cell death. *Acta Neuropathol.* 122, 75–86 (2011).
280. Gelb DJ1, Oliver E, G. S. Diagnostic Criteria for Parkinson Disease. *Arch Neurol.* 56, 33–39 (1999).
281. Bai, B. et al. U1 small nuclear ribonucleoprotein complex and RNA splicing alterations in Alzheimer's disease. (2013). doi:10.1073/pnas.1310249110/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1310249110
282. Li, Y. R., King, O. D., Shorter, J. & Gitler, a. D. Stress granules as crucibles of ALS pathogenesis. *J. Cell Biol.* 201, 361–372 (2013).
283. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25, 402–408 (2001).
284. Alieva, A. K. et al. Involvement of endocytosis and alternative splicing in the formation of the pathological process in the early stages of parkinson's disease. *Biomed Res. Int.* 2014, (2014).
285. Hamza, T. H. et al. Common genetic variation in the HLA region is associated with late-onset sporadic Parkinson's disease. *Nat. Genet.* 42, 781–785 (2010).
286. Mccarthy, J. J. *Variant.* 12, 59–64 (2012).
287. Marongiu, R. et al. Whole gene deletion and splicing mutations expand the PINK1 genotypic spectrum. *Hum. Mutat.* 28, 98 (2007).
288. Bonifati, V. et al. Mutations in the DJ-1 gene associated with autosomal recessive early-onset parkinsonism. *Science* 299, 256–259 (2003).
289. Wang, Y. et al. Evaluation of the coverage and depth of transcriptome by RNA-Seq in chickens. *BMC Bioinformatics* 12, S5 (2011).
290. Wu, P.-Y., Phan, J. H. & Wang, M. D. Assessing the impact of human genome annotation choice on RNA-seq expression estimates. *BMC Bioinformatics* 14 Suppl 1, S8 (2013).
291. Martin, I. et al. Ribosomal protein s15 phosphorylation mediates LRRK2 neurodegeneration in parkinson's disease. *Cell* 157, 472–485 (2014).
292. Kthiri, F. et al. Translational defects in a mutant deficient in YajL, the bacterial homolog of the parkinsonism-associated protein DJ-1. *J. Bacteriol.* 192, 6302–6 (2010).
293. Gehrke, S. et al. Article PINK1 and Parkin Control Localized Translation of Respiratory Chain Component mRNAs on Mitochondria Outer Membrane. *Cell Metab.* 21, 95–108 (2015).
294. Dorval, V. & Hébert, S. S. LRRK2 in transcription and translation regulation: Relevance for Parkinson's disease. *Front. Neurol.* FEB, 2–7 (2012).
295. Gehrke, S., Imai, Y., Sokol, N. & Lu, B. Pathogenic LRRK2 negatively regulates microRNA-mediated translational repression. *Nature* 466, 637–641 (2010).
296. Rhinn, H. et al. Alternative  $\alpha$ -synuclein transcript usage as a convergent mechanism in Parkinson's disease pathology. *Nat. Commun.* 3, 1084 (2012).
297. Simunovic, F. et al. Gene expression profiling of substantia nigra dopamine neurons: further insights into Parkinson's disease pathology. *Brain* 132, 1795–1809 (2009).
298. Soreq, L. et al. Deep brain stimulation induces rapidly reversible transcript changes in Parkinson's leukocytes. *J Cell Mol Med* 16, 1496–1507 (2011).
299. Chikina, M. D. et al. Low-Variance RNAs Identify Parkinson's Disease Molecular Signature in

- Blood. *Mov. Disord.* 30, n/a–n/a (2015).
300. Holcik, M. & Sonenberg, N. Translational control in stress and apoptosis. *Nat. Rev. Mol. Cell Biol.* 6, 318–327 (2005).
  301. eIF2 signaling regulates pro-inflammatory cytokine expression and bacterial invasion.
  302. Guan, B.-J. et al. Translational Control during Endoplasmic Reticulum Stress beyond Phosphorylation of the Translation Initiation Factor eIF2. *J. Biol. Chem.* 289, 12593–12611 (2014).
  303. Mounir, Z. et al. *NIH Public Access.* 4, 1–19 (2013).
  304. Teng, Y. et al. Inhibition of eIF2 $\alpha$  dephosphorylation enhances TRAIL-induced apoptosis in hepatoma cells. *Cell Death Dis.* 5, e1060 (2014).
  305. Hoozemans, J. J. M., van Haastert, E. S., Nijholt, D. a. T., Rozemuller, A. J. M. & Scheper, W. Activation of the Unfolded Protein Response Is an Early Event in Alzheimer's and Parkinson's Disease. *Neurodegener. Dis.* 10, 212–215 (2012).
  306. Bellucci, A. et al. Induction of the unfolded protein response by  $\alpha$ -synuclein in experimental models of Parkinson's disease. *J. Neurochem.* 116, 588–605 (2011).
  307. Hoozemans, J. J. M. et al. Activation of the unfolded protein response in Parkinson's disease. *Biochem. Biophys. Res. Commun.* 354, 707–11 (2007).
  308. Tettweiler, G., Miron, M., Jenkins, M., Sonenberg, N. & Lasko, P. F. Starvation and oxidative stress resistance in. *Genes Dev.* 1840–1843 (2005). doi:10.1101/gad.1311805.1840
  309. Lin, W. et al. Loss of PINK1 Attenuates HIF-1 Induction by Preventing 4E-BP1-Dependent Switch in Protein Translation under Hypoxia. *J. Neurosci.* 34, 3079–3089 (2014).
  310. Dijkstra, A. a. et al. Evidence for Immune Response, Axonal Dysfunction and Reduced Endocytosis in the Substantia Nigra in Early Stage Parkinson's Disease. *PLoS One* 10, e0128651 (2015).
  311. Mitrovich, Q. M. & Anderson, P. Unproductively spliced ribosomal protein mRNAs are natural targets of mRNA surveillance in *C. elegans*. *Genes Dev* 14, 2173–2184 (2000).
  312. Russo, A. et al. hnRNP H1 and intronic G runs in the splicing control of the human rpl3 gene. *Biochim. Biophys. Acta - Gene Regul. Mech.* 1799, 419–428 (2010).
  313. Russo, a. et al. Autoregulatory circuit of human rpl3 expression requires hnRNP H1, NPM and KHSRP. *Nucleic Acids Res.* 39, 7576–7585 (2011).
  314. Cooper-Knock, J. et al. C9ORF72 GGGGCC Expanded Repeats Produce Splicing Dysregulation which Correlates with Disease Severity in Amyotrophic Lateral Sclerosis. *PLoS One* 10, e0127376 (2015).
  315. Breig, O., Théoleyre-Schaal, O. & Baklouti, F. Combined inhibition of PI3K and activation of MAPK p38 signaling pathways trigger erythroid alternative splicing switch of 4.1R pre-mRNA in DMSO-induced erythroleukemia cells. *Cell. Signal.* 25, 2453–2461 (2013).
  316. Blaustein, M., Pelisch, F. & Srebrow, A. Signals, pathways and splicing regulation. *Int. J. Biochem. Cell Biol.* 39, 2031–2048 (2007).
  317. White, E. S. et al. Control of fibroblast fibronectin expression and alternative splicing via the PI3K/Akt/mTOR pathway. *Exp. Cell Res.* 316, 2644–2653 (2010).
  318. Zhao, S. et al. PI3K/Akt pathway mediates high glucose-induced lipid accumulation in human renal proximal tubular cells via spliced XBP-1. *J. Cell. Biochem.* 113, 3288–3298 (2012).
  319. Karni, R., Hippo, Y., Lowe, S. W. & Krainer, A. R. activates mTORC1. October 2008, 1–5 (2008).
  320. Vivarelli, S. et al. Paraquat Modulates Alternative Pre-mRNA Splicing by Modifying the

- Intracellular Distribution of SRPK2. 8, (2013).
321. Bühler, M., Steiner, S., Mohn, F., Paillusson, A. & Mühlemann, O. EJC-independent degradation of nonsense immunoglobulin-mu mRNA depends on 3' UTR length. *Nat. Struct. Mol. Biol.* 13, 462–464 (2006).
  322. Hosoda, N., Lejeune, F. & Maquat, L. E. Evidence that Poly (A) Binding Protein C1 Binds Nuclear Pre-mRNA Poly (A) Tails. *Mol. Cell. Biol.* 26, 3085–3097 (2006).
  323. Peixeiro, I. et al. Interaction of PABPC1 with the translation initiation complex is critical to the NMD resistance of AUG-proximal nonsense mutations. *Nucleic Acids Res.* 40, 1160–1173 (2012).
  324. Joncourt, R., Eberle, A. B., Rufener, S. C. & Mühlemann, O. Eukaryotic Initiation Factor 4G Suppresses Nonsense-Mediated mRNA Decay by Two Genetically Separable Mechanisms. *PLoS One* 9, e104391 (2014).
  325. Zhang, H. et al. PABPC1 interacts with AGO2, and is responsible for the microRNA mediated gene silencing in high grade hepatocellular carcinoma. *Cancer Lett.* 367, 49–57 (2015).
  326. Martin, I. et al. Ribosomal protein s15 phosphorylation mediates LRRK2 neurodegeneration in Parkinson's disease. *Cell* 157, 472–85 (2014).
  327. Febbraro, F., Giorgi, M., Caldarola, S., Loreni, F. & Romero-Ramos, M.  $\alpha$ -Synuclein expression is modulated at the translational level by iron. *Neuroreport* 23, 576–580 (2012).
  328. Zhu, J., Ding, H., Wang, X. & Lu, Q. PABPC1 exerts carcinogenesis in gastric carcinoma by targeting miR-34c. *Int. J. Clin. Exp. Pathol.* 8, 3794–802 (2015).
  329. Kabaria, S., Choi, D. C., Chaudhuri, A. D., Mouradian, M. M. & Junn, E. Inhibition of miR-34b and miR-34c enhances  $\alpha$ -synuclein expression in Parkinson's disease. *FEBS Lett.* 589, 319–325 (2015).
  330. Liu, D. et al. Cytoplasmic poly(A) binding protein 4 is highly expressed in human colorectal cancer and correlates with better prognosis. *J. Genet. Genomics* 39, 369–74 (2012).
  331. Kini, H. K., Kong, J. & Liebhaber, S. a. Cytoplasmic Poly(A) Binding Protein C4 Serves a Critical Role in Erythroid Differentiation. *Mol. Cell. Biol.* 34, 1300–1309 (2014).
  332. Nonhoff, U. et al. Ataxin-2 interacts with the DEAD/H-box RNA helicase DDX6 and interferes with P-bodies and stress granules. *Molecular biology of the cell* 18, 1385–1396 (2007).
  333. Simonsen, A. et al. The utility of  $\alpha$ -synuclein as biofluid marker in neurodegenerative diseases: a systematic review of the literature. *Biomark Med* (2015).
  334. Kim, H.-J. Alpha-Synuclein Expression in Patients with Parkinson's Disease: A Clinician's Perspective. *Exp. Neurobiol.* 22, 77–83 (2013).
  335. Hong, Z. et al. DJ-1 and alpha-synuclein in human cerebrospinal fluid as biomarkers of Parkinson's disease. *Brain* 133, 713–726 (2010).
  336. Ulla, M. et al. Is R2\* a New MRI Biomarker for the Progression of Parkinson's Disease? A Longitudinal Follow-Up. *PLoS One* 8, 1–8 (2013).
  337. Soreq, L. et al. Long non-coding RNA and alternative splicing modulations in Parkinson's leukocytes identified by RNA sequencing. *PLoS Comput. Biol.* 10, e1003517 (2014).
  338. La Cognata, V., D'Agata, V., Cavalcanti, F. & Cavallaro, S. Splicing: is there an alternative contribution to Parkinson's disease? *Neurogenetics* (2015). doi:10.1007/s10048-015-0449-x
  339. Soreq, L., Bergman, H., Israel, Z. & Soreq, H. Exon arrays reveal alternative splicing aberrations in Parkinson's disease leukocytes. *Neurodegener. Dis.* 10, 203–206 (2012).
  340. Soreq, L., Bergman, H., Israel, Z. & Soreq, H. Deep brain stimulation modulates nonsense-

mediated RNA decay in Parkinson's patients leukocytes. *BMC Genomics* 14, 478 (2013).