



**HAL**  
open science

# Identification, par criblage génomique à haut débit, de Variations du Nombre de Copies impliquées dans les formes précoces de Maladie d'Alzheimer

Solenn Legallic

► **To cite this version:**

Solenn Legallic. Identification, par criblage génomique à haut débit, de Variations du Nombre de Copies impliquées dans les formes précoces de Maladie d'Alzheimer. Génétique. 2012. hal-01472803

**HAL Id: hal-01472803**

**<https://ephe.hal.science/hal-01472803>**

Submitted on 21 Feb 2017

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

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

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

**MEMOIRE**

**Présenté**

**par**

**Solenn LE GALLIC**

**Pour l'obtention du diplôme de l'Ecole Pratique des Hautes Etudes**

**Identification, par criblage génomique à haut débit, de  
Variations du Nombre de Copies impliquées dans les  
formes précoces de Maladie d'Alzheimer**

**soutenu le 21 septembre 2012 devant le jury suivant :**

**Dr Giovanni Stevanin : Président  
Dr Anne Rovelet-Lecrux : Tuteur scientifique  
Dr Nadine Mestre-Francés : Tuteur pédagogique  
Dr Giovanni Stevanin : Rapporteur  
Dr Pascal Chambon : Examineur**

**Mémoire préparé sous la direction de :**

**Dr Anne Rovelet-Lecrux (anne.roveletlecrux@univ-rouen.fr)**

**Laboratoire de : Génétique Médicale et  
Fonctionnelle du Cancer et des Maladies  
Neuropsychiatriques-INSERMU1079**

**Directeur : Pr Thierry Frébourg**

**et de**

**Dr Nadine Mestre-Francés (nfrances@univ-montp2.fr)**

**Laboratoire de : Vieillesse Cérébral  
et Pathogenèse des Maladies  
Neurodégénératives-  
INSERM U710-EPHE-UM2**

**Directeur : Dr Jean-Michel Verdier**

# Identification, par criblage génomique à haut débit, de Variations du Nombre de Copies impliquées dans les formes précoces de Maladie d'Alzheimer

LE GALLIC Solemn

21 septembre 2012

## RESUME

L'étude des formes rares de maladie d'Alzheimer (MA) s'est avérée être une stratégie utile dans l'identification de gènes impliqués dans le déterminisme de la MA. Cette stratégie a permis d'identifier les mutations des gènes *APP*, *PSEN1* et *PSEN2* dans 77% des formes autosomiques dominantes à début précoce de MA (ADEOAD). Les variations du nombre de copies (CNVs) sont définies comme « des segments d'ADN d'au moins 1 kb pour lesquels des différences du nombre de copies ont été identifiées par comparaison entre deux génomes ». On estime à ce jour que plus de 30% des gènes sont soumis à des CNVs, et qu'un individu sur 10 porte un CNV *de novo*. Dans ce cadre, nous avons émis l'hypothèse que des variations du nombre de copies (CNVs) rares pouvaient être à l'origine de la pathologie chez certaines familles ADEOAD sans mutation identifiée dans les gènes connus, de même que chez des cas rares de MA sporadique à début précoce.

En utilisant l'hybridation génomique comparative à haute-résolution, nous avons recherché la présence des CNVs rares chez 21 cas ADEOAD non-apparentés, sans altération sur les gènes connus, et 12 cas de MA sporadique, avec un âge de début avant 55 ans.

Les analyses ont révélé la présence de 7 CNVs uniques (4 chez les ADEOAD et 3 chez les cas isolés) non retrouvés chez 1078 contrôles et 912 cas de MA à début tardif.

De façon intéressante, 4 de ces 7 réarrangements touchent des gènes (*KLK6*, *SLC30A3*, *MEOX2*, and *FPR2*) codant des protéines fortement associées aux voies métaboliques et de signalisation du peptide  $\beta$ -amyloïde.

Bien que ces CNVs soient des variants individuels rares et associés à des sous-groupes particuliers de patients, ces résultats renforcent le rôle causal, dans la MA, d'un groupe de gènes codant des molécules suspectées de longue date, de modifier le métabolisme ou la signalisation de l'A $\beta$ , et pour lesquels des modèles animaux ou cellulaires ont déjà été développés.

**MOTS-CLES** : variations du nombre de copies, maladie d'Alzheimer à début précoce, peptide A $\beta$ , CGH-array, QMPSF

## SOMMAIRE

SOMMAIRE	3
ABBREVIATIONS	5
INTRODUCTION	7
I. Maladie d'Alzheimer	7
II. Le diagnostic de la Maladie d'Alzheimer	8
A. Le diagnostic clinique	8
B. Caractérisation post-mortem de la Maladie d'Alzheimer	8

1.	Les lésions macroscopiques	8
2.	Les lésions microscopiques	9
a.	Les dégénérescences neurofibrillaires	9
b.	Les plaques amyloïdes	9
1)	Structure et composition	9
2)	Biologie de l'amyloïdogénèse	10
a)	L'Amyloid Precursor Protein et son métabolisme	10
i.	Voie non amyloïdérique	11
ii.	Voie amyloïdérique	11
	Le complexe $\gamma$ -sécrétase	11
	Le complexe $\beta$ -sécrétase	11
iii.	La toxicité d'A $\beta$	12
C.	Les biomarqueurs diagnostiques	13
1.	Les biomarqueurs du liquide céphalorachidien	13
2.	L'imagerie par résonance magnétique	13
3.	Les marqueurs scintigraphiques de plaques amyloïdes	14
D.	Génétique de la Maladie d'Alzheimer	14
1.	Les facteurs de risques génétiques et environnementaux	14
2.	Les formes Mendéliennes de la Maladie d'Alzheimer	15
a.	Les mutations du gène de l' <i>APP</i>	16
b.	Les mutations des gènes des présénilines <i>PSEN1</i> et <i>PSEN2</i>	17
c.	Les duplications du gène de l' <i>APP</i>	18
d.	La cascade amyloïde	18
III.	Les CNVs, une découverte d'une ampleur inattendue	19
A.	Application à la recherche de CNVs	20
1.	Le génotypage des CNVs	20
2.	CNVs et génome	21
3.	Pathogénicité des CNVs	23
a.	Les CNVs bénins	23
b.	Les CNVs pathogènes	23
c.	Les CNVs pathogènes avec une pénétrance incomplète et une expressivité variable	24
d.	Les CNVs de signification incertaine	25
B.	Exploration génomique	25
1.	Les outils actuels d'exploration globale du génome	25
a.	Le caryotype	25
b.	L'hybridation génomique comparative	26
1)	Les puces de CGH	26
2)	Les puces de génotypage de SNPs	28
3)	Séquençage extensif du génome	29
2.	Les outils actuels d'exploration ciblée	29
a.	Hybridation In Situ de sondes fluorescentes	29
b.	Multiplex Amplification and Probe Hybridization	30
c.	Multiplex Ligation-dependent Probe Amplification	31
d.	Quantitative Multiplex of Short Fluorescent Fragments	31
IV.	Objectifs et étapes du travail	32
A.	Criblage génomique à haut débit, des CNVs dans les formes précoces de Maladie d'Alzheimer par technique de CGH-array	32

B. Validation des remaniements précédemment identifiés et criblage de gènes candidats par technique de QMPF	33
REFERENCES BIBLIOGRAPHIQUES	34

# ABBREVIATIONS

<b>[<sup>18</sup>F]-FDDN)</b>	[ <sup>18</sup> F fluoroethyl] (méthyl amino)-2-naphtyl) éthyliidène)
<b>μM</b>	micromètre
<b>5'UTR</b>	5' untranslated region
<b>aa</b>	acide aminé
<b>AAC</b>	angiopathie amyloïde cérébrale
<b>ADEOAD</b>	<i>Autosomal Dominant Early Onset Alzheimer Disease</i>
<b>ADN</b>	acide desoxyribonucléique
<b>ADNc</b>	ADN complémentaire
<b>ADRDA</b>	<i>The Alzheimer's Disease and Related Disorders Association</i>
<b>AICD</b>	<i>APP IntraCellular Domain</i>
<b>APH1</b>	<i>anterior pharynx defective 1</i>
<b>APOE _4</b>	<i>Apolipoprotéine E allèle _4</i>
<b>APP</b>	<i>Amyloid Precursor Protein</i>
<b>ARN</b>	acide ribonucléique
<b>ARNm</b>	ARN messenger
<b>AVC</b>	accident vasculaire cérébral
<b>Aβ peptide</b>	peptide amyloïde-β
<b>BAC/PAC</b>	<i>bacterial/PI derived artificial chromosome</i>
<b>BACE-1</b>	<i>β-site APP Cleaving Enzyme 1</i>
<b>CGH</b>	hybridation génomique comparative
<b>CLU</b>	<i>clusterin</i>
<b>cM</b>	centimorgans
<b>CNPs</b>	<i>copy number polymorphisms</i>
<b>CNVR</b>	<i>CNV region</i>
<b>CNVs</b>	<i>copy number variations</i>
<b>CR1</b>	<i>complement component (3b/4b) receptor 1</i>
<b>CST3</b>	<i>cystatin 3</i>
<b>CWP</b>	plaques de type cotton wool
<b>DCOHM</b>	<i>DcoH-like protein</i>
<b>DNAJC5G</b>	<i>DnaJ (Hsp40) homolog</i>
<b>IDE</b>	<i>Insulin-degrading enzyme</i>
<b>DFT</b>	démence frontotemporale
<b>DGV</b>	<i>Database of Genomic Variants</i>
<b>DNase I</b>	<i>Deoxyribonuclease I</i>
<b>DNF</b>	dégénérescences neurofibrillaires
<b>EPHA1</b>	<i>ephrin receptor A1</i>
<b>FISH</b>	hybridation in situ fluorescente
<b>FPRI/2/3</b>	<i>formyl peptide receptor 1, 2 and 3</i>
<b>GRN</b>	<i>progranuline</i>
<b>GWAS</b>	<i>Genome-Wide Association Study</i>
<b>HapMap</b>	<i>Haplotype Map</i>
<b>HAS1</b>	<i>hyaluronan synthase 1</i>
<b>HGP</b>	<i>Human Genome Project</i>
<b>HMBS/PBGD</b>	<i>hydroxymethylbilane synthase</i>
<b>HPLC</b>	<i>High Performance Liquid Chromatography</i>
<b>IATI</b>	<i>Innotest Amyloid Tau Index</i>
<b>Indels</b>	<i>insertions and deletions</i>
<b>IPSC</b>	<i>induced pluripotent stem cells</i>
<b>IRM</b>	imagerie par résonance magnétique
<b>kb</b>	kilobase
<b>kDa</b>	kilodalton
<b>KLK6</b>	<i>kallikrein-related peptidase 6</i>

<b>LCR</b>	liquide céphalorachidien
<b>MA</b>	maladie d'Alzheimer
<b>MAGII</b>	membrane associated guanylate kinase, WW and PDZ domain
containing 1	
<b>MAP, MAPT</b>	<i>microtubule associated protein</i>
<b>MAPH</b>	<i>multiplex amplification and probe hybridization</i>
<b>Mb</b>	megabase
<b>MEOX2</b>	<i>mesenchyme homeobox 2</i>
<b>miRNA</b>	microARN
<b>MLPA</b>	<i>multiplex ligation-dependent probe amplification</i>
<b>MMSE</b>	<i>Mini Mental State Examination</i>
<b>MT</b>	microtubule
<b>NCBI</b>	<i>national center for biotechnology information</i>
<b>NCSTN</b>	<i>nicastrin</i>
<b>NEP</b>	<i>Neprilysin</i>
<b>NGS</b>	<i>Next Generation Sequencing</i>
<b>NINCDS</b>	<i>National Institute of Neurological and Communication Disorders</i>
<b>NLRP5/8</b>	<i>NLR family, pyrin domain containing 5 and 8</i>
<b>PA</b>	plaque amyloïde
<b>pb</b>	paire de base
<b>PCR</b>	<i>polymerase chain reaction</i>
<b>PEN2</b>	<i>presenilin enhancer 2 homolog</i>
<b>PHF</b>	<i>paires de filaments hélicoïdaux</i>
<b>PIB</b>	<i>Pittsburgh Compound B</i>
<b>PICALM</b>	<i>phosphatidylinositol binding clathrin assembly protein</i>
<b>PK</b>	proteinase K
<b>PMT</b>	photomultiplicateur
<b>PSEN1/2</b>	<i>présénilines 1 et 2</i>
<b>QMPSF</b>	<i>quantitative multiplex of short fluorescents fragments</i>
<b>RE</b>	reticulum endoplasmique
<b>RT</b>	<i>reverse transcriptase</i>
<b>sAPP<math>\alpha</math></b>	APP soluble $\alpha$
<b>siRNA</b>	<i>small interfering RNA</i>
<b>SLC30A3</b>	<i>solute carrier family 30 (zinc transporter), member 3</i>
<b>SNC</b>	système nerveux central
<b>SNPs</b>	<i>single nucleotide polymorphisms</i>
<b>SORL1</b>	<i>sortilin-related receptor</i>
<b>SVOP</b>	<i>SV2 related protein homolog</i>
<b>TAR</b>	thrombopénie et absence de radius
<b>TEMP</b>	tomographie par émission mono-photonique
<b>TEP</b>	tomographie par émission de positon
<b>TRIM54</b>	<i>tripartite motif containing 54</i>
<b>USP30</b>	<i>ubiquitin specific peptidase 30</i>

# Introduction

## **I. Maladie d'Alzheimer**

C'est en 1906 que le neurologue éponyme Allemand, Alois Alzheimer, va décrire pour la première fois les lésions cérébrales caractéristiques de la MA lors de l'autopsie d'une femme morte de démence à l'âge de 51 ans.

Première cause de démence neurodégénérative humaine, la maladie d'Alzheimer (MA) affecte actuellement plus de 800.000 patients en France. Elle engendre un déclin progressif des facultés cognitives, se répercutant sur le comportement et l'adaptation sociale des patients. Peu à peu, une destruction des cellules nerveuses se produit dans les régions du cerveau liées à la mémoire et au langage. Avec le temps, la personne atteinte a de plus en plus de difficultés à mémoriser les événements, à reconnaître les objets et les visages, à se rappeler la signification des mots et à exercer son jugement. Cette pathologie présente un fort impact sur l'autonomie et conduit généralement au décès entre 3 et 15 ans après le diagnostic. À ce jour, la MA est toujours une maladie incurable dont on ne sait que ralentir l'évolution. L'âge en est le principal facteur de risque. En effet, l'incidence de la maladie double tous les 5 ans après 65 ans (Hirtz et al., 2007). Elle ne constitue toutefois pas une pathologie spécifique du sujet âgé.

On distingue actuellement les formes sporadiques représentant la majorité des patients et pour lesquelles le facteur de risque génétique majeur est le génotype Apolipoprotéine E allèle \_4 (APOE \_4) ainsi que d'autres facteurs moindres (Lambert et al., 2009) et les formes répondant à un déterminisme mendélien. Si il reste pour le moment difficile d'appréhender pleinement la physiopathologie de la MA sporadique, la découverte de plusieurs anomalies génétiques impliquées dans les formes à transmission mendélienne a permis d'établir certaines hypothèses physiopathologiques, elles-mêmes à l'origine de nouvelles thérapeutiques en cours de validation.

## **II.**



## Le diagnostic de la Maladie d'Alzheimer

### . Le diagnostic clinique

En 1983, un groupe d'experts réunis au sein du National Institute of Neurological and Communication Disorders (NINCDS) and The Alzheimer's Disease and Related Disorders Association (ADRDA) a établi des recommandations pour le diagnostic clinique de la MA (McKhann et al., 1984).

Les critères de ce rapport communément appelé critères du NINCDS-ADRDA permettent de définir les éléments cliniques nécessaires au diagnostic de MA probable ou possible (Lopez et al., 1990). Ce diagnostic basé sur la réalisation de tests d'évaluation neuropsychologique standardisés, comme le *Mini Mental State Examination* (MMSE), s'appuie sur l'existence de troubles cognitifs affectant au moins deux domaines cognitifs dont l'un doit être mnésique. Par ailleurs, il doit exister une perte d'autonomie au moment du diagnostic et la durée d'évolution doit être supérieure à 6 mois. Si l'ensemble des critères est présent, le diagnostic de MA probable peut être retenu.

Ces critères servent encore de nos jours aussi bien dans le domaine de la recherche pour d'innombrables études que dans la pratique clinique quotidienne des neurologues, gériatres et psychiatres. Cependant, cet ensemble de critères souffre de plusieurs limites, la principale étant une faible spécificité, d'environ 70% (Knopman et al., 2001), par comparaison avec d'autres démences. C'est pourquoi, toujours dans le cadre des critères NINCDSADRDA, seule l'étude anatomopathologique du cerveau, chez le sujet *postmortem*, permet de poser le diagnostic de MA avec certitude, bien qu'il soit désormais possible de compléter le diagnostic de la MA par l'usage de biomarqueurs les plus spécifiques et les plus précoces possibles.

## 0. Caractérisation post-mortem de la Maladie d'Alzheimer

### 1. Les lésions macroscopiques

L'examen macroscopique du cerveau montre une atrophie cérébrale, qui s'accompagne d'une dilatation des ventricules cérébraux, d'un élargissement des sillons corticaux, ainsi que d'une perte neuronale diffuse. (Whitehouse et al., 1981).

Cependant, cet examen reste peu contributif, le vieillissement s'accompagnant d'une atrophie et d'une perte de poids du cerveau dont le degré et l'âge d'apparition sont controversés (Rapoport, 1990) et un examen microscopique reste indispensable.

## 2. Les lésions microscopiques

Les deux lésions histologiques caractéristiques de la MA observées au début du siècle dernier sont les plaques amyloïdes (PA) et les dégénérescences neurofibrillaires (DNF).

Elles sont observables en microscopie optique après coloration adaptée (Perl, 2000).

### a. **Les dégénérescences neurofibrillaires**

Les DNF constituent un des deux types de lésions histopathologiques caractéristiques de la MA. Les DNF sont des lésions intraneuronales dues à l'accumulation, dans le cytoplasme et les dendrites des neurones en dégénérescence, de structures filamenteuses sous forme de paires de filaments hélicoïdaux (*Paired Helical Filaments*, PHF) (Kidd, 1963; Terry, 1963).

Le composant principal de ces DNF est une protéine du cytosquelette appartenant à la famille des *microtubule associated protein* (MAP), la protéine Tau, dont la fonction est la stabilisation et la polymérisation des microtubules (MT). Dans les DNF, Tau est hyperphosphorylée, détachée des MT et va former les PHF (Brion et al., 1985).

Il est possible de visualiser ces DNF sur les coupes histologiques par des anticorps spécifiques après imprégnation argentique d'agrégats visibles en microscopie optique (méthodes de Gallyas, Bielschowsky ou Bodian) (Duyckaerts et al., 1990; Vallet et al., 1992).

On peut rencontrer des DNF au cours du vieillissement physiologique (Polvikoski et al., 2010) cependant leur localisation se limite alors aux régions hippocampiques et elles n'atteignent que de manière succincte les aires associatives postérieures comme c'est le cas dans la MA. Enfin, les DNF permettent de graduer le stade de la maladie. Il existe en effet une bonne corrélation entre l'importance de ces lésions et le déclin cognitif (Braak and Braak, 1991 ; Delacourte et al., 1999).

### b. **Les plaques amyloïdes**

#### 1) **Structure et composition**

Les PA sont des lésions extracellulaires apparaissant comme des masses sphériques, d'un diamètre de 5 à 100 micromètres ( $\mu\text{M}$ ), et résultant de l'accumulation d'un fragment peptidique insoluble appelé peptide amyloïde- $\beta$  ( $A\beta$ ) de 4 kilodaltons (kDa).

Ce peptide est issu du clivage protéolytique de son précurseur, la protéine

Amyloid Precursor Protein (APP) (Glennner and Wong, 1984). Le caractère amyloïde de ce peptide vient de sa capacité à s'agréger sous forme de feuillets  $\beta$  plissés, qui peuvent être colorés par des substances comme la thioflavine S et le rouge Congo.

L'utilisation d'anticorps spécifiques dirigés contre ce peptide A<sub>n</sub> et chacune de ses isoformes, a permis de distinguer les différents types de plaques. Il existe différents types de plaques dont la quantité et la répartition évoluent progressivement au cours de la maladie. Enfin, d'autres constituants mineurs ont également été identifiés au sein de ces plaques, tels l'APOE, l'antichymotrypsine, le facteur du complément C1, l' $\alpha$ 2-macroglobuline et un fragment de l' $\alpha$ -synucléine (Abraham et al., 1988; Schmechel et al., 1993; Ueda et al., 1993).

## 2) **Biologie de l'amyloïdogénèse**

La découverte du principal composant des PA, le peptide A $\beta$ , a naturellement placé celui-ci au centre du processus pathologique. Dès lors, la recherche de la protéine, et donc du gène dont est issu ce peptide, s'est avérée essentielle. La purification et le séquençage du peptide A $\beta$  (Glennner et Wong, 1984) a permis d'identifier un acide desoxyribonucléique (ADN) complémentaire codant pour l'APP (Tanzi et al., 1987).

### a) **L'Amyloid Precursor Protein et son métabolisme**

Le gène codant l'APP est situé sur le bras long du chromosome 21, en 21q21.3. L'APP est une protéine transmembranaire de 770 aa (acides aminés), ubiquitaire, dont l'expression est particulièrement élevée au niveau du système nerveux central (SNC) aussi bien dans les neurones que les cellules gliales (Selkoe, 1994). Son rôle propre n'est pas parfaitement établi mais elle semble intervenir dans la plasticité neuronale et la formation synaptique.

Durant la voie sécrétoire, l'APP mature est soumis à plusieurs coupures enzymatiques. Le métabolisme de l'APP emprunte deux voies distinctes, dites amyloïdergique et non amyloïdergique, selon que le peptide A $\beta$  est produit ou non. Ces deux voies se distinguent essentiellement l'une de l'autre par le type d'enzymes impliquées dans le clivage de l'APP et par les métabolites qui en résultent. Le clivage protéolytique d'APP se produit sur des sites bien spécifiques et fait appel à 3 protéines ou complexes distincts ; l'\_\_, la \_\_ et la \_\_-sécrétase.

**i. Voie non amyloïdérique**

La voie non amyloïdérique est la plus commune, puisqu'elle intervient dans tous les types cellulaires. Elle fait intervenir deux clivages endoprotéolytiques orchestrés par l' $\alpha$ -sécrétase, puis la  $\gamma$ -sécrétase, et engendre trois fragments distincts : la partie APP soluble  $\alpha$  (sAPP $\alpha$ ) libérée dans l'espace extracellulaire, le domaine *APP IntraCellular Domain* (AICD), et le peptide P3. Cette voie de maturation de l'APP correspond à un processus physiologique et ne produit pas de peptide A $\beta$ . On considère actuellement que le peptide sAPP $\alpha$  aurait une activité neuroprotectrice et un rôle utile dans la plasticité neuronale (Thornton et al., 2006).

**ii. Voie amyloïdérique**

*Le complexe  $\beta$ -sécrétase*

Dans la voie amyloïdérique, l'APP est clivé par la  $\beta$ -sécrétase au lieu de l' $\alpha$ -sécrétase sur un autre site. L'enzyme responsable,  *$\beta$ -site APP Cleaving Enzyme 1* (BACE-1), (Lin et al., 2000; Vassar et al., 1999; Yan et al., 1999), est une protéine glycosylée transmembranaire de type 1 et principalement localisée dans l'appareil de Golgi et les endosomes (Capell et al., 2000; Vassar et al., 1999).

Le clivage par BACE-1 libère d'une part un domaine extracellulaire, l'APP soluble  $\beta$  et un fragment C-terminal de 99 aa, le C99.

*Le complexe  $\gamma$ -sécrétase*

A la suite du clivage par la  $\beta$ -sécrétase, l'APP peut subir, à l'extrémité C-terminale de la séquence A $\beta$ , une maturation par le complexe  $\gamma$ -sécrétase libérant d'une part l'AICD intracellulaire et d'autre part le peptide A $\beta$ . Le clivage au niveau de l'extrémité C-terminale présente une certaine hétérogénéité produisant des peptides A $\beta$  de longueur variable (39 à 45 acides aminés) (Selkoe, 1994) parmi lesquels les isoformes 40 (A $\beta$ <sub>40</sub>) et 42 (A $\beta$ <sub>42</sub>), indiquant respectivement la longueur d'acides aminés, sont les plus représentées. L'A $\beta$ <sub>42</sub> est considéré comme l'isoforme la plus neurotoxique en raison de sa propriété hautement insoluble et oligomérisante (Jarrett et al., 1993).

Quatre protéines participent à la formation du complexe protéique  $\gamma$ -sécrétase, parmi lesquelles les présénilines 1 et 2 (PSEN1 et PSEN2) jouent un rôle capital dans le processus de clivage d'APP (De Strooper et al., 1998).

Chaque PSEN est codée par un gène spécifique : respectivement *PSEN1* situé sur le bras long du chromosome 14 et *PSEN2* situé sur le bras long du chromosome 1. Les PSEN sont largement exprimées dans le SNC (Sherrington et al., 1995).

La seconde protéine identifiée dans le complexe  $\gamma$ -sécrétase est la nicastrine (NCSTN) (Yu et al., 2000), une glycoprotéine transmembranaire de type I exprimée préférentiellement dans le cerveau. Au niveau cellulaire, elle est localisée dans le reticulum endoplasmique (RE) et l'appareil de Golgi. Des travaux ont montré que l'extrémité N-terminale du domaine transmembranaire de la NCSTN est impliquée dans l'interaction avec le complexe  $\gamma$ -sécrétase (Capell et al., 2003; Shah et al., 2005).

Enfin, peu de temps après la découverte de la NCSTN, les derniers composants du complexe, *anterior pharynx defective 1 (APH1)* et *presenilin enhancer 2 homolog (PEN2)*, furent isolés au cours d'un crible génétique dans le nématode recherchant des mutants de l'activité  $\gamma$ -sécrétase (Francis et al., 2002; Goutte et al., 2002). APH1 présente trois isoformes dont la principale, APH1B, est exprimée dans le cerveau et semble être un facteur de stabilisation des PSEN. L'orthologue PEN2 humain code une protéine à deux domaines transmembranaires qui serait requise pour le clivage endoprotéolytique des PSEN (Luo et al., 2003; Shiraishi et al., 2004; Takasugi et al., 2003).

La formation du complexe dépendant des PSEN est régie par une séquence d'événements maintenant bien caractérisée. La NCSTN forme tout d'abord un complexe avec APH1, puis ce sous-complexe interagit avec la PSEN1 ou la PSEN2, et enfin PEN2 s'associe au complexe pour promouvoir vraisemblablement l'hydrolyse des PSEN, conduisant au complexe biologiquement actif (Takasugi et al., 2003).

### iii. La toxicité d'A<sub>n</sub>

En condition physiologique, le peptide A<sub>42</sub> est minoritaire par rapport au peptide A<sub>40</sub> dans les conditions normales, représentant entre 5 et 10% de la production totale.

Les propriétés biophysiques des peptides 40 et 42 sont très différentes, l'isoforme 42 présentant une capacité d'agrégation bien plus élevée (Finder and Glockshuber, 2007). L'accumulation intraneuronale de peptide a été notée dans les cerveaux de patients Alzheimer à des stades très précoces avant l'apparition des plaques extracellulaires (Cataldo et al., 2004). L'agrégation du peptide A<sub>n</sub> commence probablement au sein des neurones avant que le peptide A<sub>n</sub> soit secondairement libéré dans les espaces extracellulaires, conduisant à la formation des plaques amyloïdes (Langui et al., 2004). L'agrégation de la protéine A<sub>n</sub> commence par la formation d'oligomères puis s'organise

en amas de taille plus conséquente. Ces oligomères ont une toxicité propre avec un effet délétère sur la potentialisation à long terme hippocampique et la densité synaptique dès les stades précoces de la maladie (Shankar et al., 2007).

## Les biomarqueurs diagnostiques

### 1. Les biomarqueurs du liquide céphalorachidien

A partir des lésions anatomopathologiques et de leurs constituants biochimiques, plusieurs études ont identifié des biomarqueurs biologiques spécifiques de la MA présents dans le liquide céphalorachidien (LCR). Les biomarqueurs actuellement les plus pertinents sont le peptide A<sub>42</sub>, constituant des plaques amyloïdes, la protéine Tau tout isoforme confondu, dite alors totale (t-Tau ou Tau) et la protéine Tau anormalement phosphorylée en position 181 (P-tau ou Phospho-Tau) identifiée dans les DNF.

Dans la MA, la concentration d'A<sub>42</sub> du LCR diminue alors que la quantité de Tau et P-Tau augmente comparativement aux témoins (Vandermeeren et al., 1993; Motter et al., 1995; Hu et al., 2002). Une solution pour augmenter la spécificité consiste à utiliser différents rapports entre ces biomarqueurs comme le IATI (*Innotest Amyloid Tau Index*) ou le rapport A<sub>42</sub>/Phospho-Tau (de Souza et al., 2011).

### 2. L'imagerie par résonance magnétique

Par définition, un biomarqueur idéal doit être le plus sensible et spécifique possible pour une maladie donnée, peu invasif, et reproductible. Le développement des techniques d'imagerie par résonance magnétique (IRM), a permis d'établir les marqueurs iconographiques de lésions cérébrales.

L'atrophie temporale interne notamment au niveau hippocampique et parahippocampique semble être « très fréquente » (71-96%) dans la MA et « fréquente » dans les troubles cognitifs légers (59-78%) alors qu'elle reste « limitée » dans la population de sujets sains (29%) (Visser et al., 2005; De Leon et al., 1997). De plus, elle apparaît comme étant corrélée à la présence et à la sévérité des lésions anatomopathologiques de la MA (Jack et al., 2002). Cette atrophie a donc été proposée comme un des nouveaux critères de diagnostic en recherche (Dubois et al., 2007).

### 3. Les marqueurs scintigraphiques de plaques amyloïdes

Les technologies de tomographie par émission de positon (TEP) et par émission mono-photonique (TEMP) permettent d'étudier *in vivo* par des radiotraceurs spécifiques

le débit sanguin cérébral et le métabolisme du glucose, qui est le reflet du débit sanguin ou du métabolisme cérébral. Ces radiotraceurs, disponibles en pratique clinique courante, apportent une aide diagnostique. On utilise également ces techniques, d'une excellente sensibilité, pour le suivi évolutif de la maladie.

Depuis quelques années, ont été développés des radiotraceurs capables de lier de manière spécifique les PA sous forme fibrillaire et/ou les DNF. Trois radio-ligands semblent apporter un réel gain sur le plan diagnostique. Le *Pittsburgh Compound B (PIB)*, le [ $^{18}\text{F}$  fluoroethyl] (méthyl amino]-2-naphtyl) éthylidène) ([ $^{18}\text{F}$ ]-FDDNP) et l'AV-45 constituent des molécules de choix pour l'avenir (Feldman et al., 2011) même si actuellement elles ne sont disponibles que dans un cadre de recherche, compte-tenu des limites techniques pour les produire et des coûts importants.

## Génétique de la Maladie d'Alzheimer

### 1. Les facteurs de risques génétiques et environnementaux

Dans le cas des formes sporadiques de la maladie, dont l'âge de début est classiquement plus tardif que les formes autosomiques dominantes, différents facteurs de risques génétiques sont connus (Lambert et al., 2009) dont le principal est l'allèle  $\epsilon_4$  du gène APOE (APOE4) codant pour l'apolipoprotéine E.

Ainsi 10% des porteurs d'un allèle APOE4 ayant atteint l'âge de 75 ans développeront une MA et 33% si ils sont homozygotes APOE4/E4 (Genin et al., 2011). L'importance de ce facteur de risque est attestée par de nombreuses études épidémiologiques (Bertram and Tanzi, 2008), ou d'études d'association tout-génome (*Genome-Wide Association Study*, GWAS) (Harold et al., 2009; Lambert et al., 2009). Plusieurs hypothèses sont avancées pour expliquer le lien physiopathologique entre l'isoforme APOE4 et l'apparition de lésions de la maladie. Il semblerait que cette protéine interagisse avec le peptide A $\beta$  en modifiant ses capacités d'agrégation et sa clairance mais elle pourrait également intervenir dans la plasticité neuronale ou la neuroinflammation (Kim et al., 2009).

Récemment deux études en GWA ont identifié des polymorphismes sur trois autres gènes (Harold et al., 2009; Lambert et al., 2009): *clusterin (CLU)*, *phosphatidylinositol binding clathrin assembly protein (PICALM)* et *complement component (3b/4b) receptor 1 (CR1)* conférant un risque indépendant respectivement de 0,86 pour les 2 premiers et de 1,21 pour le dernier. Ces données ont rapidement été confirmées par d'autres études et méta-analyses (Carrasquillo et al., 2010; Jun et al., 2010). Les

protéines codées par *CLU* et *CRI* seraient indirectement impliquées dans les processus d'élimination du peptide A<sub>β</sub> (Holtzman, 2004; Rogers et al., 2006). Le rôle de ces facteurs de risque semble donc être lié aux mécanismes d'agrégation et d'élimination du peptide A<sub>β</sub>.

D'autre part, les études épidémiologiques ont permis de définir également différents facteurs de risque épigénétiques et environnementaux. L'âge et par conséquent le vieillissement est le facteur de risque principal de la MA. A celui-ci s'ajoutent plusieurs autres facteurs comme les pathologies vasculaires (Kudo et al., 2000), les traumatismes crâniens (Jellinger, 2004), l'alcool et le tabac.

## 2. Les formes Mendéliennes de la Maladie d'Alzheimer

L'immense majorité des cas de MA répond à un déterminisme multifactoriel et correspond dans ce cas soit à des formes sporadiques sans antécédents familiaux, soit à des formes familiales à transmission non mendélienne.

Dans une minorité de cas, la MA répond à un déterminisme mendélien à transmission autosomique dominante. L'ADEOAD est un phénomène rare avec une prévalence de 5,3 pour 100.000 personnes, soit environ 1000 cas en France et se caractérise par un début précoce de la maladie avant l'âge de 60 ans (Campion et al., 1999).

Ces formes présentent une diversité phénotypique clinique, la démence pouvant être associée à une paraparésie spastique, un syndrome extrapyramidal précoce, un syndrome cérébelleux ou à un accident vasculaire cérébral (AVC) résultant d'une angiopathie amyloïde sévère. Il existe également une diversité neuropathologique par la possible association de plaques de type cotton wool (CWP), de nombreux corps et neurites de Lewy, ou d'une angiopathie amyloïde.

L'étude de ces formes autosomiques dominantes est décisive pour la compréhension physiopathologique de la maladie d'Alzheimer puisque l'altération des gènes impliqués est suffisante pour provoquer la maladie.

L'identification de ces causes génétiques a permis de formuler l'hypothèse amyloïdergique (Hardy et Higgins, 1992) qui est à l'origine des développements thérapeutiques actuels concernant le peptide A. Actuellement trois gènes sont identifiés dans ces formes (*APP*, *PSEN1*, *PSEN2*) et leurs mutations sont impliquées dans 77% des 144 familles françaises, originaires de l'ensemble du territoire, étudiées depuis 1993 au laboratoire (Campion et al., 1999 ; Dumanchin et al., 2006 ; Wallon et al.,



2012).

**a. Les mutations du gène de l'APP**

Dès 1987, l'étude de familles dans lesquelles la MA ségrégeait de façon autosomique dominante a permis de mettre en évidence une liaison génétique entre les loci 21q21 et 21q22-1 avec la pathologie (St George-Hyslop et al., 1987). En 1991, une région de 21 centimorgans (cM) a pu être clairement associée à la maladie. Cette région contenant le gène *APP*, celui-ci a été préférentiellement séquencé, permettant de mettre en évidence la première mutation impliquée dans une forme autosomique dominante (Goate et al., 1991).

A ce jour, 33 mutations ponctuelles du gène *APP* affectant 90 familles ont été décrites comme responsables de formes familiales de la MA (<http://www.molgen.vib-ua.be/ADMutations/default.cfm?MT=1&ML=6&Page=StatPerGene>).

L'âge d'apparition de la maladie dans ces familles est compris entre 30 et 65 ans (Rogaeva et al., 2006).

La plupart des mutations du gène *APP* sont des mutations faux-sens localisées dans les exons 16 et 17. On distingue deux effets distincts de ces mutations en fonction de leur localisation au niveau des sites de coupure  $\beta$  et  $\gamma$  sécrétase, ou au sein même de la séquence du peptide  $A\beta$ .

Ces mutations agissent sur l'agrégation du peptide  $A_{40}$  de trois manières différentes : (i) la production excessive de fragments  $A_{40}$  et  $A_{42}$  pour les mutations concernant le site de clivage  $\beta$ -sécrétase, (ii) la modification du rapport de concentration entre ces 2 peptides pour les mutations du site de clivage  $\alpha$ -sécrétase, (iii) augmentation de l'agrégabilité pour les mutations de la région codante, (Hardy, 1997; Wolfe, 2007). Dans ces cas, ainsi que pour la mutation Swedish (Lys595Asp/Met596Leu) au site  $\beta$ , la MA peut être associée à une *angiopathie* amyloïde cérébrale (AAC) plus ou moins sévère, notamment avec certaines mutations Flemish (Ala692Gly), Arctic (Glu693Gly), Italian (Glu693Lys) et Iowa (Asp694Asn).

Plus récemment, une équipe Italienne a décrit la première mutation récessive du gène *APP* à effet dominant négatif sur l'amyloïdogénèse (Di Fede et al., 2009).

**b.**

## Les mutations des gènes des présénilines *PSEN1* et *PSEN2*

En 1992, une étude rapporta une liaison génétique très significative entre des marqueurs du chromosome 14 et la MA au sein de certaines familles atteintes de formes précoces (Schellenberg et al., 1992). En 1994 a été décrite l'existence de mutations pathogènes sur un gène alors inconnu, qui fut dénommé *PSEN1* (Sherrington et al., 1995).

Le gène *PSEN1* est considéré comme le gène majoritairement responsable des formes autosomiques dominantes précoces de MA. A ce jour, 185 mutations différentes affectant 405 familles ont été décrites pour ce gène, indiquant une considérable hétérogénéité allélique (<http://www.molgen.vib-ua.be/ADMutations/default.cfm?MT=1&ML=6&Page=StatPerGene>) Il s'agit pour la plupart de mutations ponctuelles de type faux-sens, et plus rarement de mutations d'épissage au niveau de l'intron 8, entraînant une délétion de l'exon 9.

Il est actuellement admis que la conséquence biologique de ces mutations est une modification des propriétés de la préséniline favorisant la coupure A $\beta$ <sub>42</sub> aux dépens de la coupure A $\beta$ <sub>40</sub>, à l'origine d'une augmentation du rapport A<sub>42</sub>/A<sub>40</sub> dans le sang, le liquide céphalorachidien et le cerveau. La mesure des niveaux d'expression d'A<sub>40</sub> et A<sub>42</sub> *in vitro* dans des cellules transfectées avec différentes mutations *PSEN1* (Kumar-Singh et al., 2006) a montré une augmentation de A<sub>42</sub>/A<sub>40</sub>, liée à la surproduction d'A<sub>42</sub> dans la moitié des cas et à une diminution de A<sub>40</sub> pour quelques mutations.

Dans la grande majorité des cas, les mutations de *PSEN1* sont responsables d'une MA sans signe associé. Cependant, il existe dans plusieurs familles une très grande diversité phénotypique liée aux mutations de *PSEN1*, associant aux troubles cognitifs une paraparésie spastique (Karlstrom et al., 2008), d'importantes modifications comportementales (Zekanowski et al., 2006) ou une ataxie cérébelleuse (Anheim et al., 2007). Certaines mutations de *PSEN1* sont ainsi associées à des CWP qui consistent en de larges dépôts, principalement constitués de peptide A $\beta$ <sub>42</sub> (Steiner et al., 2001; Tabira et al., 2002).

De façon générale, la pénétrance de la MA liée à des mutations de *PSEN1* est complète avant 60 ans mais il existe une diversité des âges de début inter-mutations. Dans une étude *in vitro* mesurant la production d'A<sub>40</sub>, A<sub>42</sub> et le ratio A<sub>42</sub>/A<sub>40</sub> selon les différentes mutations de *PSEN1*, il a été montré une corrélation significative entre l'âge de début de MA et l'augmentation du ratio A<sub>42</sub>/A<sub>40</sub> (Kumar-Singh et al., 2006).

Parallèlement à la découverte du gène *PSEN1*, une étude de liaison génétique

sur plusieurs familles d'origine germanique présentant une forme monogénique familiale de MA excluait l'implication des chromosomes 21 et 14. Pour ces familles, une liaison significative fut obtenue sur le chromosome 1, impliquant un locus situé en 1q31-42 (Levy-Lahad et al., 1995). La mise en évidence d'une très forte homologie de séquence entre un ADN complémentaire de ce locus et le gène *PSENI* a alors conduit à la mise en évidence d'un troisième gène pathogène, appelé *PSEN2*.

A ce jour, seules 13 mutations faux-sens de ce gène ont été décrites chez 22 familles (<http://www.molgen.vib-ua.be/ADMutations/default.cfm?MT=1&ML=6&Page=StatPerGene>). Ces mutations, contrairement à celles de *PSENI*, sont associées à un âge de début de la maladie plus tardif, entre 40 et 85 ans, et une pénétrance parfois incomplète (Rogaev et al., 1995; Sherrington et al., 1996).

#### **c. Les duplications du gène de l'*APP***

Enfin, des duplications du locus du gène *APP* ont été identifiées dans une vingtaine de familles (Cabrejo et al., 2006 ; Rovelet-Lecrux et al., 2006 ; Sleegers et al., 2006 ; Guyant- Marechal et al., 2008 ; Kasuga et al., 2009 ; McNaughton et al., 2010).

Cette anomalie quantitative est responsable d'une surexpression de la protéine *APP* qui entraîne elle-même l'augmentation de production de peptide A<sub>β</sub> (Pottier et al., 2011).

Les duplications sont de taille variable allant de 0,58 megabases (Mb) à 6,37Mb, mais le phénotype reste indépendant de la taille de la duplication car dans une famille la duplication ne concernait que l'*APP* excluant l'implication potentielle des gènes adjacents (Sleegers et al., 2006). De même, l'*APP* étant localisé sur le chromosome 21, sa duplication suffit à expliquer que les patients trisomique 21 ont un risque accru de développer une MA après l'âge de 40 ans (Engidawork and Lubec, 2001).

#### **d. La cascade amyloïde**

La compréhension des effets biologiques de ces diverses mutations a permis une avancée majeure dans la caractérisation de la physiopathologie de la maladie. En effet, il est apparu que les mutations de l'*APP* ou des *PSENI/2*, exprimées dans des modèles cellulaires avaient pour conséquence commune une augmentation du ratio A<sub>β</sub>42/A<sub>β</sub>40 secrété, les formes A<sub>β</sub>42 représentant les formes les plus agrégables du peptide A<sub>β</sub>. Ce rapport augmenté a également été retrouvé dans les cerveaux des

patients, dans leurs liquides biologiques et chez les animaux transgéniques exprimant ces gènes mutants.

Ces données ont permis de formuler l'hypothèse de la cascade amyloïdérique (Hardy, 2007) qui stipule que toutes ces mutations interfèrent avec la maturation du peptide A $\beta$  et que, dans ces formes, l'accumulation anormale de formes agrégées du peptide A est le *primum movens* de la maladie d'Alzheimer.

### **III. Les CNVs, une découverte d'une ampleur inattendue**

Les observations cytogénétiques ont détecté depuis longtemps la présence de larges duplications et délétions dans le génome humain. Mais leur fréquence était estimée faible et leur présence habituellement reliée à des maladies génétiques rares. La première observation reconnue pour être associée à une maladie fut la présence d'un chromosome 21 additionnel avec le syndrome de Down (Lejeune et al., 1959).

L'avènement de nouvelles technologies associées à des stratégies bioinformatiques a conduit à une meilleure résolution de détection des anomalies chromosomiques et à la caractérisation de nombreuses régions variables, répétées ou polymorphes du génome impliquant des segments de taille inférieure à celles vues en microscopie (3Mb), mais de taille supérieure à celles vues par une technique de séquençage (1Kb).

Une avancée considérable de ces dernières années dans l'étude du génome humain a été le HGP qui a abouti en 2001 à la publication de la séquence complète du génome humain haploïde, permettant la découverte de nombreux gènes et stimulant le lancement de nouveaux projets pour étudier la structure du génome humain.

Parmi ceux-ci, le projet « *Haplotype Map (HapMap)* » débuté en 2002 et terminé en 2009 qui avait pour objectif de développer une carte des haplotypes de l'ensemble du génome humain à l'aide de millions de SNPs répartis sur l'ensemble du génome.

En 2004, deux études de CGH (de type 85K et BACs, respectivement) réalisées sur un total de 75 individus « normaux » ont révélé un nombre inattendu de ces remaniements (Itsara et al., 2009; Sebat et al., 2004).

La meilleure résolution de détection des anomalies chromosomiques a conduit à caractériser de nombreux remaniements de taille variable appelés « variants de structure », dont l'effet pathogène n'est pas toujours prouvé. Cependant leur contribution aux modifications du génome humain apparaît aussi importante que celle des variations nucléotidiques détectées par séquençage, les SNPs.

Certains variants n'ont aucune conséquence phénotypique et d'autres

influencent le dosage génique et peuvent entraîner un phénotype, seuls ou combinés à d'autres facteurs environnementaux, ils sont alors nommés « anomalies de structure ». Leur fréquence est beaucoup plus grande que les remaniements microscopiques et leur implication dans les pathologies semble beaucoup plus importante.

Leur cartographie dans le génome à partir de 270 individus HapMap a permis d'avoir une idée plus précise de l'ampleur et des caractéristiques de ces remaniements (Redon et al., 2006). Ces variations génomiques furent désignées sous le terme de variations du nombre de copies ou *copy number variations* (CNVs) et définies comme « des segments d'ADN d'au moins 1kb présentant un nombre de copies variable, par comparaison entre deux génomes » (Feuk et al., 2006). Ces variations structurales non équilibrées peuvent être des délétions, des duplications (en tandem ou insérées à un autre endroit du génome) ou des amplifications d'un plus grand nombre de copies.

Les gains ou pertes de copies d'une taille inférieure à 1kb sont appelés « indels » (*insertions and deletions*).

Le terme CNV est un terme général qui n'implique aucune donnée de fréquence ou d'effet phénotypique. Les CNVs polymorphes, qui sont retrouvés dans plus de 1% de la population générale, sont appelés *copy number polymorphisms* (CNP). Certains loci comportent plusieurs CNVs proches ou se recouvrant, ce type de région est appelé *CNV region* (CNVR).

Le projet « 1000 génomes », initié en 2008 et toujours en cours, s'inscrit dans la lignée du projet HapMap avec, en plus des SNPs, l'inclusion dans la carte génomique des indels et des CNPs.

Enfin, les GWAS sont des études d'associations pangénomiques entre haplotypes et maladies, utilisant les données générées par les projets « HapMap » et « 1000 génomes ». Elles sont un outil puissant pour aider à l'identification de gènes de susceptibilité ou de séquences modulatrices de gènes dont les mutations sont à l'origine de maladies monogéniques.

## Application à la recherche de CNVs

### 1. Le génotypage des CNVs

Les études de population ont recensé des milliers de CNVs dont la plupart ont une taille supérieure à 5kb due aux limites de résolution des techniques actuelles. Ces études se sont cependant majoritairement concentrées sur la

découverte de nouveaux CNVs et non sur leur génotypage.

Une étude récente sans précédent (Conrad et al., 2010) a combiné découverte et génotypage de CNVs communs (>1%) supérieurs à 1kb. Dans un premier temps, les auteurs ont mis en évidence des CNVs supérieurs à 500 pb chez 40 individus européens et africains à l'aide de 20 puces de CGH comportant 2,1 millions de sondes oligonucléotidiques chacune soit 42 millions au total réparties sur la portion testable du génome avec un espacement médian de 56 pb.

Ils ont ainsi mis en évidence 11700 loci de CNVs ayant une taille médiane de 2,7kb dont 49% n'ont été identifiés que chez un individu.

10819 de ces CNVs observés (92%) ont ensuite été analysés sur une puce de génotypage de 105000 sondes oligonucléotidiques chez 450 individus HapMap (européens, japonais et africains) et 45 individus chinois. 5238 CNVs non redondants ont ainsi pu être génotypés. Parmi ceux ci, 77% étaient des délétions, 16% des duplications et 7% étaient multi alléliques. Le ratio 5:1 des délétions sur les duplications reflète probablement en partie le challenge technique plus important du génotypage des duplications.

## 2. CNVs et génome

La comparaison de deux génomes dans cette même étude (Conrad et al., 2010) révèle, à ce degré de résolution, une moyenne de 1098 CNVs par individu et une taille cumulée des loci de CNVs de 24Mb soit 0.78% du génome.

Les 8599 CNVs validés chez les 41 individus (40 et l'individu de référence) couvrent un total de 112.7Mb, soit 3.7% du génome.

Il est intéressant de noter qu'en comparant 2 génomes, 40,5% des CNVs touchent 3,1% des gènes RefSeq, altérant la structure de 2,7% des transcrits et altérant directement la séquence codante de 1,2% des ARN messagers (ARNm).

Si tous les CNVs de cette étude sont considérés collectivement, 38,8% des CNVs touchent 13,4% des gènes RefSeq, altérant la structure de 12,5% des transcrits et altérant la séquence codante de 5,5 des ARNm.

Plus de la moitié des délétions partielles qui touchent des exons sont prédites comme induisant des décalages du cadre de lecture. Les délétions emportant des gènes entiers provoquent la perte de fonctions de 267 gènes. Un appauvrissement

des délétions et des CNVs communs chevauchant des gènes a été observé par rapport aux duplications ou aux loci multi-alléliques et aux CNVs rares. Ce biais des délétions communes hors des gènes est plus accentué chez les africains que chez les européens, en accord avec la plus faible sélection contre les substitutions de bases délétères chez les européens (Lohmueller et al., 2008).

Les CNVs sont également retrouvés hors des enhanceurs et des éléments ultraconservés mais pas hors des promoteurs ni des sites sensibles à la Deoxyribonuclease I (DNase I). Les duplications semblent être enrichies au niveau des promoteurs et des codons stop. Les indels avaient par ailleurs été décrits comme enrichis aux extrémités des gènes (Ng et al., 2008). Ng et collaborateurs suggèrent que les variants aux extrémités des gènes seraient probablement neutres, notamment parce qu'il existe des codons start et stop alternatifs pouvant produire une protéine fonctionnelle. Ils ne seraient donc pas contre-sélectionnés.

Les gènes impliqués dans les processus biologiques extracellulaires tels que l'adhésion, la reconnaissance et la communication cellulaires ainsi que les gènes impliqués dans la perception sensorielle et le système immunitaire sont enrichis dans les CNVRs. A l'inverse, les gènes impliqués dans les processus intracellulaires tels que les voies de biosynthèse et métabolique et dans les complexes multi protéiques sont sous représentés dans les CNVRs. Une étude préalable avait déjà montré que les CNVs touchaient préférentiellement des gènes codant pour des protéines situées à la périphérie des réseaux protéiques en interaction et des cellules (partie extracellulaire ou membranaire), phénomène probablement dû à une augmentation des événements adaptatifs en réponse à un changement d'environnement (Kim et al., 2007).

La *Database of Genomic Variants (DGV)* (<http://dgvbeta.tcag.ca/dgv/app/home?ref=NCBI36/hg18>) recensait, au 2 novembre 2010, 66741 CNVs à 15963 loci et 34229 indels, résultat d'une augmentation exponentielle du nombre de publications depuis sa création en 2004.

Une phase plateau est constatée depuis 2010, expliquée par l'avènement de l'aire du séquençage tout-génome.

De manière intéressante, les statistiques de cette base de données permettent d'observer une corrélation inverse entre la taille des CNVs et leur fréquence. On peut noter une surreprésentation des CNVs de 100 à 200kb qui reflète le biais

introduit par les puces de type BAC puisque ces études, largement utilisées dans cette base de données, rapportent les bornes du clone BAC lui-même, même s'il s'agit en réalité de variants plus petits comme nous l'avons vu. Il y a donc un décalage artificiel vers la taille standard des BACs qui est habituellement de 150-250kb.

En ce qui concerne les indels, on observe une forte augmentation des variants de 100 à 500pb reflétant probablement l'importance des séquences Alu dont c'est la taille dans le génome humain.

### 3. Pathogénicité des CNVs

Dans une revue parue en 2012, V. Malan et S. Romana traitent du problème d'interprétation que posent les CNVs. Selon eux, « Les moyens techniques ont fait des progrès plus rapides que notre compréhension des implications biologiques et médicales de l'information générée. La signification d'un CNV identifié dans la pathogénèse du phénotype observé peut parfois être très difficile à établir. Certains CNVs sont clairement pathogènes mais parfois avec une pénétrance incomplète et une expressivité variable alors que d'autres posent des problèmes d'interprétation. D'autres encore sont considérés comme bénins. » (Malan & Romanan, 2012)

#### a. **Les CNVs bénins**

Jusqu'à récemment, il était communément admis que l'ADN entre deux individus était identique à 99,9%. Seules les variations « qualitatives » de l'ADN liées à des modifications de bases nucléotidiques étaient alors connues.

L'étude génomique de témoins par CGH-*array* a mis en évidence l'existence de CNVs considérés comme bénins (Redon et al., 2006). L'étude de ces CNVs a montré qu'il s'agissait le plus souvent de répétitions en tandem de cassettes d'ADN de taille variable (la plupart d'entre elles ont une taille inférieure à 100kb) pouvant comprendre des séquences codantes et non codantes. Le nombre de répétition de ces cassettes (de 0 à plus de 20) varie d'un individu à l'autre et définit les allèles du CNV qui se transmettent de manière mendélienne.

Lorsque ces CNVs sont retrouvées chez plus de 1% de la population, on parle de polymorphismes ou *Copy Number Polymorphisms* ou CNPs. S'ils sont présents dans moins de 1% de la population, on parle de « variants bénins privés ».

La majorité de ces CNPs sont sans conséquence phénotypique actuellement décelée alors que d'autres sont associés à des maladies multifactorielles fréquentes



(comme le lupus) ou sont le témoin de l'adaptation humaine à l'environnement (Zhang et al., 2009 ; Sudmant et al., 2010).

**b. Les CNVs pathogènes**

Malan et Romana définissent qu'un CNV est pathogène « lorsqu'il est responsable des signes cliniques observés chez le patient. Ces déséquilibres génomiques pathogènes surviennent le plus souvent *de novo* et ont une taille généralement supérieure à 400 kb. Par ailleurs, le phénotype est d'autant plus sévère que la taille du déséquilibre chromosomique est grande et comprend de nombreux gènes. On admet qu'à partir de 1,5 Mb, un CNV est quasiment toujours à l'origine de la pathologie observée (Cooper et al., 2011). »

**c. Les CNVs pathogènes avec une pénétrance incomplète et une expressivité variable**

Selon Malan et Romana, « Certains CNV identifiés chez des patients avec un phénotype anormal sont nécessaires à la survenue de la pathologie mais ils ne sont pas suffisants en raison de leur présence chez un des deux parents sains ou des individus normaux. Ces CNVs rares confèrent un risque plus ou moins important de développer la pathologie. Un exemple éloquent où il a été mis en évidence un CNV rare avec une pénétrance incomplète correspond au syndrome TAR (Thrombopénie et Absence de Radius). Cette pathologie associe une thrombopénie centrale et une absence bilatérale de radius. On note une grande variabilité phénotypique incluant, à des degrés divers, des anomalies cardiaques, gastro-intestinales, squelettiques et hématologiques. Par une approche CGH-*array*, une équipe Allemande a mis en évidence une délétion 1q21 d'environ 200 kb chez tous les patients présentant un syndrome TAR (Klopocki et al., 2007). Cette délétion n'a pas été retrouvée chez 700 individus contrôles et n'est pas répertoriée dans la DGV. Cependant, environ 75% des parents sont porteurs de l'anomalie sans être atteints. Des facteurs génétiques ainsi que des facteurs épigénétiques et environnementaux peuvent modifier la pénétrance et l'expressivité d'une pathologie génétique. Dans le syndrome TAR, la présence d'un ou de plusieurs gène(s) modificateur(s) pourrait expliquer la pénétrance incomplète et l'expressivité variable de la maladie. »

Par ailleurs, l'utilisation de la CGH a révélé des déséquilibres génomiques associés à des pathologies psychiatriques et neurologiques. Par exemple, des délétions de 300 kb à 1Mb comprenant le locus du gène *CHRNA7*, localisé en 15q13, peuvent

être associées à des épilepsies isolées. Dans ces pathologies, il existe une très grande variabilité phénotypique intra et inter familiale et une pénétrance incomplète. C'est ainsi que les délétions du gène *CHRNA7*, peuvent être responsables hormis l'épilepsie, de troubles psychiatriques (autisme, schizophrénie, troubles bipolaires), ou de déficience intellectuelle (Guilmatre et al., 2009 ; Sharp et al., 2008). Par ailleurs, des délétions en 16p11.2 d'une taille de 200 kb sont associées à une obésité alors que des duplications de cette même région sont associées à des retards de croissance et à des formes d'anorexie car elles impliquent le gène *SHB2* qui joue un rôle clé dans le contrôle cérébral des signaux qui régissent le stockage des graisses, l'utilisation de sucre, le bilan énergétique et le poids (Walters et al., 2010 ; Jacquemont et al., 2011). Ainsi, le champ de pathologies associées aux CNVs devient de plus en plus large.

#### **d. Les CNVs de signification incertaine**

Environ 10% des CNV soulèvent des difficultés d'interprétation. Ils correspondent pour la plupart à des duplications (70% des cas) d'une taille d'environ 700kb et héritées de l'un des deux parents sains ou retrouvés dans des populations témoins (Cooper et al., 2011).

Plusieurs éléments sont à considérer pour déterminer si le CNV identifié est délétère ou non. Une table avec différents critères d'évaluation des CNV pour aider à l'interprétation des résultats a ainsi été proposée (Miller et al., 2010). Les arguments majeurs pour affirmer qu'un CNV est pathogène sont : son caractère *de novo*, sa taille supérieure à 400 kb, son contenu riche en gènes et l'existence d'autres patients avec le même CNV dont le phénotype est anormal (il peut s'agir de syndromes connus ou non). A l'inverse, un CNV hérité d'un parent sain, décrit chez des individus normaux dans les bases de données, pauvre en gènes et de petite taille est très souvent bénin.

### Exploration génomique

#### **1. Les outils actuels d'exploration globale du génome**

##### **a. Le caryotype**

Les premières observations de chromosomes datent de 1880 par Flemming mais ce n'est qu'en 1956 que le nombre de chromosomes a été établi à 46 chez l'homme par Tjio et Levan qui ont eu l'idée, à partir de cultures de tissus, d'introduire le choc hypotonique pour améliorer l'étalement des chromosomes.

Actuellement, le caryotype classique diffère peu de cette époque (Caspersson et al., 1968). Les cellules, souvent des lymphocytes, sont mises en culture pendant trois

jours avant d'être bloquées en métaphase de mitose en présence de colchicine qui perturbe les fuseaux mitotiques, puis gonflées dans un milieu hypotonique et fixées sur lames de verre. Cet étalement chromosomique est ensuite dénaturé, afin d'altérer certaines régions chromosomiques, soit à la chaleur, soit par technique enzymatique, et est enfin coloré au Giemsa faisant apparaître des bandes sombres et claires sur les chromosomes. La topographie des bandes est caractéristique d'un chromosome.

En 1959, Jérôme Lejeune et collaborateurs décrivent ainsi la première anomalie chromosomique associée à une pathologie, la trisomie 21 (Lejeune et al., 1959). La même année, Jacobs et Strong décrivent la première anomalie des chromosomes sexuels, avec une formule XXY, dans le syndrome de Klinefelter (Jacobs and Strong, 1959). En 1960, la première nomenclature de classification des chromosomes est décrite, basée sur leur taille et la position de leur centromère. Les causes de syndromes malformatifs chromosomiques ont ensuite été découvertes très rapidement, tels que le syndrome de Turner ou les trisomies 13 et 18. En 1960, Nowell et Hungerford (Nowell and Hungerford, 1960) décrivent également la première anomalie chromosomique associée à une affection maligne, le chromosome de Philadelphie, qui se révélera être une translocation  $t(9;22)$  par la suite.

Le caryotype, encore utilisé en première intention à l'heure actuelle, ne permet d'observer que des anomalies du nombre de chromosomes et des remaniements de plus de 5Mb visibles en microscopie. Les remaniements plus petits n'étaient donc pas étudiés à l'époque où seule cette technique existait.

## **b. L'hybridation génomique comparative**

### **1) Les puces de CGH**

Afin de palier à ces limites, l'*hybridation génomique comparative* (*Comparative Genomic Hybridization*, CGH) a été développée en 1992 (Kallioniemi et al., 1992).

Le principe est basé sur l'hybridation simultanée et compétitive d'un ADN génomique test et d'un ADN génomique témoin en présence d'ADN Cot-I, permettant de saturer les régions répétées. Les deux ADN sont digérés et marqués par des haptènes ou des fluorophores différents. Ces ADN étaient à l'origine hybridés en quantité équimolaires sur des métaphases normales et les séquences spécifiques d'un chromosome homologue présentes dans ces deux échantillons entraient alors en compétition pour la même cible

chromosomique. La différence d'intensité de fluorescence (directe ou révélée par des anticorps couplés à un fluorochrome) entre les deux échantillons permettait alors de déterminer les régions amplifiées ou délétées dans l'échantillon test. Cette technique permettait une analyse pan génomique mais à une résolution très faible, de l'ordre de 10Mb.

Le remplacement des chromosomes métaphasiques par des puces oligonucléotidiques (Lipshutz et al., 1999), rendues possibles par les avancées techniques et les données du HPG, a engendré l'explosion de cette technique. Les puces désormais utilisées sont de petites lames de verre (ou de silicone) sur lesquelles sont immobilisées des sondes. Chaque sonde est une séquence d'ADN spécifique d'une séquence donnée du génome humain. Elles sont organisées en spots sur la puce, chacun contenant un grand nombre de copies d'une même sonde. Les puces sont fabriquées par dépôts de *bacterial/PI derived artificial chromosome* (BAC/PAC) ou par synthèse *in situ* des sondes par jet d'encre (Agilent®) ou photolithographie (Affymetrix®). Les spots de sondes sont organisés de façon bien définie afin qu'un scanner puisse lire la lame après hybridation et annoter les spots et ainsi déterminer les régions présentant des remaniements non équilibrés. Le principe de la technique est resté le même qu'en 1992, cependant l'utilisation de puces a permis une amélioration de la résolution qui ne cesse d'augmenter. La résolution dépend en effet de la taille des séquences déposées sur la lame, du nombre et de l'espacement des sondes au niveau génomique.

Ainsi, les premières puces utilisaient des sondes de type cosmides ou des BACs de plusieurs dizaines de kilobases (kb) (Pinkel et al., 1998). Ces puces ont l'avantage de proposer une hybridation robuste et une bonne couverture, avec un bon rapport signal sur bruit, mais ne permettent pas la détection de remaniements inférieurs à 50 kb pour les BACs ou à 15kb pour les cosmides. De plus, les BACs et cosmides étant caractérisés par des tailles importantes, les remaniements mis en évidence avec ces puces présentent des tailles aberrantes qui ne peuvent descendre en dessous de la taille du BAC alors que des remaniements bien plus petits peuvent interférer avec leur hybridation.

Les sondes peuvent aussi être des clones d'ADN complémentaire (ADNc) ou des produits de *polymerase chain reaction* (PCR) génomiques variant d'une centaine à

plusieurs milliers de paires de bases (pb).

Les puces les plus récentes et les plus résolutes mais avec un moins bon rapport signal sur bruit, utilisent ainsi des sondes oligonucléotidiques de 25 à 80pb. Les premières puces de CGH présentaient 44000 sondes, elles atteignent maintenant les 2,7 millions de sondes espacées d'1kb environ permettant une très haute résolution de moins de 5kb, d'un facteur mille par rapport aux techniques de cytogénétique classique. Cette résolution peut encore être augmentée par la possibilité de réaliser des puces à façon permettant l'enrichissement en sondes dans des régions d'intérêt. Il est préconisé de considérer la variation de trois à cinq sondes successives suivant la résolution de la lame utilisée pour valider un remaniement.

Comme toute technique, elle présente des limites, notamment le manque de couverture des régions répétées du génome qui en représentent une fraction non négligeable et l'absence de détection de remaniements équilibrés.

2)

### Les puces de génotypage de SNPs

Cette technologie de puce à ADN a également été utilisée pour des puces de génotypage ou puces à *single nucleotide polymorphisms* (SNPs), qui permettent de génotyper un grand nombre de SNPs sur une même puce et servent aussi à la détection de remaniements génomiques. Il existe deux principes différents selon le distributeur.

Affymetrix® propose la digestion de l'ADN génomique suivie de la ligature d'adaptateurs. Les fragments sont alors amplifiés à l'aide d'amorces spécifiques des adaptateurs. Les produits de PCR sont purifiés, fragmentés, marqués à une extrémité et hybridés sur la puce à SNPs. Une sonde spécifique à chaque allèle est présente sur la lame.

Les différents allèles vont donc s'hybrider sur la sonde correspondante, permettant un génotypage par apparition d'un signal s'il y a « *perfect match* » et d'une absence de signal s'il y a « *mismatch* ». Il peut exister des biais d'amplification intragénomique et interindividuelle.

Illumina® propose une amplification de l'ADN génomique suivie d'une fragmentation puis d'une hybridation sur puce. La technologie Illumina® repose sur la fixation des sondes à des billes de silicone (technologie *BeadArray*). Chaque bille présente une sonde en des milliers d'exemplaires et chaque bille est représentée trente fois. Une extension d'amorce est réalisée après hybridation des fragments. La base ajoutée par extension complémentaire au SNP est reliée à un haptène qui sera reconnu par un anticorps couplé à un fluorophore différent en fonction de la base, permettant de discriminer les allèles et ainsi de réaliser le génotypage par scan.

La comparaison des intensités de fluorescence entre individus permet également d'estimer le nombre relatif de copies du locus associé à un SNP donné. Le calcul du nombre de copies est généralement réalisé sur plusieurs sondes de SNPs du même segment génomique, rendant le résultat d'autant plus fiable. L'avantage de cette technique est sa haute résolution qui a notamment permis de constater que les premiers remaniements décrits par puce de type BAC (Redon et al., 2006) présentaient des tailles réelles généralement 5 à 15 fois plus petites (McCarroll et al., 2008). Ces données indiquent que le nombre de gènes et la proportion du génome touchés par des variations du nombre de copies est probablement bien inférieure aux premières descriptions. Les

bornes estimées par puces de SNPs diffèrent des points de cassure déterminés par séquençage d'une médiane de 1,6kb seulement (McCarroll et al., 2008).

Cependant, il existe des problèmes majeurs inhérents à cette technique, souvent ignorés ou compensés par un éventail d'algorithmes. Ces puces, largement utilisées, présentent des limites importantes en termes de sensibilité et de couverture voire de spécificité. L'interprétation de leurs données doit donc être faite avec prudence.

### 3) **Séquençage extensif du génome**

Avec les récentes avancées dans le domaine de la génomique, les techniques de séquençage haut débit (*Next Generation Sequencing*, NGS) se sont développées rapidement, permettant le lancement d'études pangénomiques. Cette technique permet le séquençage de millions de fragments d'ADN simultanément.

Cependant, pour des raisons de coût et d'interprétation de la masse de données générées, le séquençage se concentre souvent sur la partie du génome codant pour les protéines, l'exome. Il ne représente qu'environ 1,5% du génome, mais du fait de son caractère codant, l'exome regroupe près de 85% des mutations ayant un effet majeur sur des phénotypes pathologiques (Majewski et al., 2011). En effet, la plupart des maladies mendéliennes sont causées par des mutations exoniques ou des sites d'épissage canoniques, qui altèrent la séquence en acides aminés du gène affecté. Le séquençage d'exomes est une avancée majeure qui s'impose aujourd'hui dans la plupart des cas pour l'identification des causes des pathologies mendéliennes, qui résultent largement de mutations exoniques ou de mutations de sites d'épissage (Lalonde et al., 2010 ; Ng et al., 2010).

Le séquençage à haut débit permet de détecter des variations de plus petite taille que ceux détectés par les plateformes CGH notamment les indels inférieurs à 70 pb. Le séquençage permet l'identification exacte de la localisation des bordures des remaniements, lorsque ceux-ci sont localisés au sein d'un exon. Malgré tout, la CGH reste une solution plus économique pour l'étude d'un génome en entier. De plus, le séquençage haut-débit ne règle pas le problème de détection des CNVs dans les régions répétées, qui sont mal capturées et par la suite difficiles à aligner sur le génome de référence.

## 2. **Les outils actuels d'exploration ciblée**

### a. **Hybridation In Situ de sondes fluorescentes**

Le développement des techniques d'hybridation in situ fluorescente (*fluorescent in*

*situ hybridization*, FISH) et de nombreuses sondes spécifiques de régions chromosomiques a révolutionné la cytogénétique classique, avec une résolution moyenne de 150 à 200 kb.

Des sondes d'ADN (ou d'acide ribonucléique (ARN)), marquées par un fluorochrome ou par un haptène puis un anticorps couplé à un fluorochrome, et ciblant une région génomique connue, sont mises en contact avec les chromosomes sur noyaux interphasiques ou chromosomes en métaphase et vont aller s'hybrider sur leur séquence cible si celle-ci est présente. Ces sondes sont ensuite visualisées par microscopie à fluorescence. Le principe d'hybridation *in situ* date des années 1970 mais utilisait la radioactivité et était donc beaucoup plus lourd à appliquer.

La FISH est maintenant utilisée en routine dans les laboratoires de cytogénétique en complément du caryotype et offre de multiples possibilités. Il existe ainsi différents types de sondes : des sondes de peinture chromosomique spécifiques d'un chromosome entier, d'un bras voire de quelques bandes chromosomiques, des sondes locus spécifiques correspondant à un petit fragment chromosomique et des sondes télomériques et centromériques.

La FISH permet de détecter des anomalies du nombre de chromosomes homogènes ou en mosaïque, d'identifier un fragment chromosomique dans le cadre de la détection de matériel supplémentaire d'origine inconnue sur un chromosome ou dans des remaniements complexes et de mettre en évidence des microdélétions ou microduplications non visibles sur le caryotype standard, et ce de manière rapide et précise. Cependant, malgré la possibilité d'un marquage multicolore permettant l'analyse de plusieurs régions simultanément, la FISH reste une technique ciblée nécessitant une orientation cytogénétique ou clinique préalable.

D'autres techniques de biologie moléculaire se sont aussi développées telles que le génotypage de microsatellites pour rechercher une perte d'hétérozygotie ou le Southern Blot pour repérer une éventuelle délétion ou insertion en comparant les tailles des fragments reconnus par une sonde les ciblant spécifiquement, mais elles présentent les mêmes limites que la FISH.

#### **b. Multiplex Amplification and Probe Hybridization**

La *multiplex amplification and probe hybridization* (MAPH) ( Armour et al., 2000)



nécessite une première étape d'hybridation de l'ADN génomique sur une membrane de nylon. Des sondes correspondant aux séquences cibles (jusqu'à 40) y sont ensuite hybridées. Ces sondes, de tailles différentes afin de pouvoir les discriminer par électrophorèse, présentent des séquences universelles en 5' et 3' permettant par la suite leur amplification par PCR. Après hybridation, les sondes non liées sont éliminées par lavage, les sondes restantes étant alors représentatives du nombre de copies des séquences cibles. Ces sondes sont ensuite déshybridées puis amplifiées à l'aide d'amorces universelles, et les produits de PCR sont séparés par électrophorèse puis quantifiés. Cette technique relativement fastidieuse a été remplacée par la MLPA.

### **c. Multiplex Ligation-dependent Probe Amplification**

La *multiplex ligation-dependent probe amplification* (MLPA) a été développée par Schouten et collaborateurs en 2002 (Schouten et al., 2002). Deux oligonucléotides adjacents sont utilisés comme sondes spécifiques de chaque région. Les extrémités 5' et 3' contiennent des séquences universelles permettant l'utilisation d'amorces universelles de PCR. Pour différencier les fragments amplifiés à partir de chaque région, une séquence variable est ajoutée entre la sonde et la séquence universelle en 3'. Après hybridation à leur cible, les sondes sont liguées en présence d'une ADN ligase thermostable. Les produits de ligature sont ensuite amplifiés par PCR en conditions quantitatives avec un couple d'amorces universelles, dont l'une est marquée par un fluorophore. Enfin, les produits de PCR sont séparés selon leur taille par électrophorèse, puis quantifiés et normalisés. La quantité relative de produits d'amplification correspond au nombre relatif de copies de la séquence cible. Cette technique, qui permet de détecter des délétions et duplications sur 40 cibles différentes simultanément, est applicable à la caractérisation de SNPs, dans ce cas les sondes diffèrent d'un nucléotide spécifique du SNP au niveau du point de ligature. Une puce MLPA a également été décrite permettant l'analyse d'un plus grand nombre de cibles (Zeng et al., 2008).

### **d. Quantitative Multiplex of Short Fluorescent Fragments**

Enfin, la *quantitative multiplex of short fluorescent fragments* (QMPSF) est basée sur l'amplification simultanée de plusieurs petites séquences d'ADN génomique à l'aide de couples d'amorces dont l'une est couplée à un fluorochrome.

L'amplification est réalisée dans un nombre limité de cycles de PCR afin de se placer dans la phase exponentielle et de rester ainsi quantitative. Les produits sont ensuite séparés en fonction de leur taille sur un séquenceur automatique. Les électrophérogrammes du patient et d'un témoin sont superposés informatiquement et ajustés au niveau d'un amplicon contrôle correspondant à un gène non impliqué dans la pathologie étudiée. Une délétion hétérozygote sera visualisée par une diminution de moitié de la hauteur du pic de fluorescence et une duplication hétérozygote par une augmentation de 50% de la hauteur du pic (Casilli et al., 2002; Charbonnier et al., 2000; Saugier Veber et al., 2001).

Cette technique, à la fois simple à mettre en œuvre et particulièrement flexible permet ainsi l'analyse simultanée de plusieurs cibles (une vingtaine) sur de grandes cohortes à faible coût et en peu de temps. Développée au sein de l'Unité Inserm 1079, la technique de QMPSF a fait l'objet de nombreuses applications pour la détection des altérations quantitatives dont la mise en évidence d'une duplication du gène *APP* comme cause majeure de maladie d'Alzheimer associée à une angiopathie amyloïde cérébrale (Rovelet-Lecrux et al., 2006). Un des risques associés à cette technique reste l'existence de faux positifs. En effet, des variations nucléotidiques situées dans les régions génomiques correspondant aux zones d'amorces aboutissent à un non amorçage de l'amorce et à une diminution de moitié du pic correspondant, mimant une délétion. Les délétions doivent donc être confirmées par séquençage des zones d'amorces de l'amplicon concerné.

#### **IV. Objectifs et étapes du travail**

##### **Criblage génomique à haut débit, des CNVs dans les formes précoces de Maladie d'Alzheimer par technique de CGH-array**

La recherche de mutations au sein de 144 familles ADEOAD recrutées en France au sein de notre laboratoire a démontré que les mutations sur les gènes *APP*, *PS1*, *PS2* ainsi que les duplications du gène *APP* expliquaient 77% des cas d'ADEOAD (Wallon et al., 2012).

Par ailleurs, nous avons également démontré que certains CNVs peuvent être à l'origine de pathologies à transmission mendélienne, comme dans le cas de la maladie d'Alzheimer, dont certaines formes résultent d'une duplication du gène *APP*, ainsi que dans les démences frontotemporales (DFT), avec les délétions des gènes *progranuline*

(*GRN*) (Rovelet-Lecrux et al., 2008) et *microtubule associated protein (Tau)* (Rovelet-Lecrux et al., 2009).

Nous avons donc supposé que de rares CNVs pouvaient être responsables de la MA dans les familles ADEOAD chez qui l'étude des gènes connus s'est avérée négative.

Par ailleurs, la fréquence importante des remaniements survenant *de novo* fait des CNVs rares pathogènes un mécanisme d'altération génétique envisageable dans des cas sporadiques.

C'est pourquoi nous avons décidé d'étudier de la même façon un groupe de cas de MA sporadique à début précoce en supposant la présence de CNVs *de novo* dans ce groupe.

Afin de tester cette hypothèse, nous avons mené sur ces familles, une recherche de CNVs tout-génome au moyen d'une plate-forme de CGH-*array* installée au laboratoire et dotée d'un scanner de troisième génération, le *DNA Microarray Scanner with SureScan High Resolution Technology* (Agilent Technologies, France) sur un format de lames 1x1000k afin de déterminer les bornes de ces remaniements le plus précisément possible, excepté pour un patient pour qui un format 4x180k a été utilisé.

### Validation des remaniements précédemment identifiés et CRIBLAGE de gènes candidats par technique de QMPSF

Notre but étant d'identifier des CNVs rares causaux, exclusivement associés à ces deux groupes de MA, on ne s'attend pas à les retrouver dans un groupe de 1078 contrôles.

Enfin, pour renforcer la spécificité de nos résultats, propres à ces phénotypes particuliers de MA, nous avons également recherché la présence de CNVs associés de façon exclusive avec ces deux groupes de MA dans une troisième cohorte de 912 MA à début tardif. De précédentes études sur cette population avaient recherché en vain des CNVs (Heinzen et al., 2010 ; Swaminathan et al., 2011).

La pathogénicité des CNVs retenus a été établie sur un faisceau d'arguments : absence dans une large population témoin, co-ségrégation avec la maladie dans les formes familiales, contenu en gènes des régions remaniées.

Afin de valider sur une deuxième technique indépendante les remaniements identifiés par CGH, nous avons modélisé au fur et à mesure des résultats obtenus, 3 QMPSF contenant des amplicons choisis sur ces régions mais également sur des

gènes d'intérêt que nous souhaitons tester sur nos cohortes. Parmi ces gènes d'intérêt, en premier lieu, les gènes candidats directement impliqués dans la voie amyloïde dont *NCSTN*, *APH1* et *PEN2* du complexe  $\gamma$ -sécrétase (Marlow et al., 2003 ; Carrasquillo et al., 2010) et *BACE-1* du complexe  $\beta$ -sécrétase. Enfin, *LRP* (Jaeger et al., 2009), *Neprilysin (NEP)* (Marr et al., 2004) et *Insulin-degrading enzyme (IDE)* (Vepsäläinen et al., 2007 ; 2008) sont connus pour agir sur la clairance de l'A<sub>β</sub>.

D'autre part, nous avons choisi de tester les « tops gènes » issus des GWAS récentes dont les principaux sont *CLU*, *PICALM*, *CRI* (Harold et al., 2009 ; Lambert et al., 2009) (Carrasquillo et al., 2010 ; Corneveaux et al., 2010 ; Jun et al., 2010) et *ephrin receptor A1 (EPHA1)* (Morgan K., 2009). La *cystatin 3 (CST3)* et *sortilin-related receptor (SORL1)* (Rogaeva et al., 2007 ; Reitz et al., 2011) ont été identifiés de la même façon.

## Références bibliographiques

Abraham, C. R., Selkoe, D. J. & Potter, H. Immunochemical identification of the serine protease inhibitor alpha 1-antichymotrypsin in the brain amyloid deposits of Alzheimer's disease. *Cell* 52, (1988).

Anheim, M. et al. Ataxic variant of Alzheimer's disease caused by Pro17Ala PSEN1 mutation. *Journal of neurology, neurosurgery, and psychiatry* 78, (2007).

Annas, G. J. Privacy rules for DNA databanks. Protecting coded 'future diaries'. *JAMA : the journal of the American Medical Association* 270, (1993).

Armour, J. A., Sismani, C., Patsalis, P. C. & Cross, G. Measurement of locus copy number by hybridisation with amplifiable probes. *Nucleic acids research* 28, (2000).

Bertram, L. & Tanzi, R. E. Thirty years of Alzheimer's disease genetics: the implications of systematic meta-analyses. *Nature reviews. Neuroscience* 9, (2008).

Braak, H. & Braak, E. Neuropathological staging of Alzheimer-related changes. *Acta neuropathologica* 82, (1991).

Brion, J. P., Couck, A. M., Passareiro, E. & Flament-Durand, J. Neurofibrillary tangles of Alzheimer's disease: an immunohistochemical study. *Journal of submicroscopic cytology* 17, (1985).

Buxbaum, J.D. et al. Molecular dissection of NRG1-ERBB4 signaling implicates PTPRZ1 as a potential schizophrenia susceptibility gene. *Mol Psychiatry* 13, (2008).

Cabrejo, L. et al. Phenotype associated with APP duplication in five families. *Brain : a journal of neurology* 129, (2006).

Campion, D. et al. Early-onset autosomal dominant Alzheimer disease: prevalence, genetic heterogeneity, and mutation spectrum. *American journal of human genetics* 65,

(1999).

Capell, A. et al. Maturation and pro-peptide cleavage of beta-secretase. *The Journal of biological chemistry* 275, (2000).

Carrasquillo, M. M. et al. Concordant association of insulin degrading enzyme gene (IDE) variants with IDE mRNA, Abeta, and Alzheimer's disease. *PloS one* 5, e8764 (2010).

Carrasquillo, M. M. et al. Replication of CLU, CR1, and PICALM associations with alzheimer disease. *Archives of neurology* 67, (2010).

Casilli, F. et al. Rapid detection of novel BRCA1 rearrangements in high-risk breast-ovarian cancer families using multiplex PCR of short fluorescent fragments. *Human mutation* 20, (2002).

Caspersson, T. et al. Chemical differentiation along metaphase chromosomes. *Experimental cell research* 49, (1968).

Cataldo, A. M. et al. Abeta localization in abnormal endosomes: association with earliest Abeta elevations in AD and Down syndrome. *Neurobiology of aging* 25, (2004).

Charbonnier, F. et al. Detection of exon deletions and duplications of the mismatch repair genes in hereditary nonpolyposis colorectal cancer families using multiplex polymerase chain reaction of short fluorescent fragments. *Cancer research* 60, (2000).

Chaudhury, A.R. et al. Neuregulin-1 and erbB4 immunoreactivity is associated with neuritic plaques in Alzheimer disease brain and in a transgenic model of Alzheimer disease. *J Neuropathol Exp Neurol* 62, (2003).

Cirulli, E. T. & Goldstein, D. B. Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nature reviews. Genetics* 11, (2010).

Conrad, D. F. et al. Origins and functional impact of copy number variation in the human genome. *Nature* 464, (2010).

Cooper, G.M. et al. A copy number variation morbidity map of developmental delay. *Nat Genet* 43 (2011).

Corneveaux, J. J. et al. Association of CR1, CLU and PICALM with Alzheimer's disease in a cohort of clinically characterized and neuropathologically verified individuals. *Human molecular genetics* 19, (2010).

de Leon, M. J. et al. Frequency of hippocampal formation atrophy in normal aging and Alzheimer's disease. *Neurobiology of aging* 18, (1997).

de Souza, L. C. et al. Cerebrospinal fluid biomarkers in the differential diagnosis of Alzheimer's disease from other cortical dementias. *Journal of neurology, neurosurgery, and psychiatry* 82, (2011).

de Strooper, B. et al. Deficiency of presenilin-1 inhibits the normal cleavage of amyloid precursor protein. *Nature* 391, (1998).

Delacourte, A. et al. The biochemical pathway of neurofibrillary degeneration in aging and Alzheimer's disease. *Neurology* 52, (1999).

Deshpande, A., Kawai, H., Metherate, R., Glabe, C. G. & Busciglio, J. A role for synaptic zinc in activity-dependent Abeta oligomer formation and accumulation at excitatory synapses. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 29, (2009).

Di Fede, G. et al. A recessive mutation in the APP gene with dominant-negative effect on amyloidogenesis. *Science* 323, (2009).

Dubois, B. et al. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. *Lancet neurology* 6, (2007).

Dubois, B. et al. Revising the definition of Alzheimer's disease: a new lexicon. *Lancet neurology* 9, (2010).

Duclos, A. et al. Pitfalls in the use of DGV for CNV interpretation. *Am J Med Genet A* 155A, (2011).

Dumanchin, C. et al. Segregation of a missense mutation in the microtubule-associated protein tau gene with familial frontotemporal dementia and parkinsonism. *Hum Mol Genet.* 7, (1998).

Dumanchin, C. et al. Biological effects of four PSEN1 gene mutations causing Alzheimer disease with spastic paraparesis and cotton wool plaques. *Human mutation* 27, (2006).

Duyckaerts, C. et al. Rating of the lesions in senile dementia of the Alzheimer type: concordance between laboratories. A European multicenter study under the auspices of EURAGE. *Journal of the neurological sciences* 97, (1990).

Engidawork, E. & Lubec, G. Protein expression in Down syndrome brain. *Amino acids* 21, (2001).

Feuk, L., Carson, A. R. & Scherer, S. W. Structural variation in the human genome. *Nature reviews. Genetics* 7, (2006).

Finder, V. H. & Glockshuber, R. Amyloid-beta aggregation. *Neuro-degenerative diseases* 4, (2007).

Francis, R. et al. aph-1 and pen-2 are required for Notch pathway signaling, gamma-secretase cleavage of betaAPP, and presenilin protein accumulation. *Developmental cell* 3, (2002).

Friedlich, A. L. et al. Neuronal zinc exchange with the blood vessel wall promotes cerebral amyloid angiopathy in an animal model of Alzheimer's disease. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 24, (2004).

Genin, E. et al. APOE and Alzheimer disease: a major gene with semi-dominant inheritance. *Molecular psychiatry* 16, (2011).

George-Hyslop, P. H. et al. The genetic defect causing familial Alzheimer's disease maps on chromosome 21. *Science (New York, N.Y.)* 235, (1987).

Glenner, G. G. & Wong, C. W. Alzheimer's disease: initial report of the purification and characterization of a novel cerebrovascular amyloid protein. *Biochemical and biophysical research communications* 120, (1984).

Goate, A. et al. Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. *Nature* 349, (1991).

Goutte, C., Tsunozaki, M., Hale, V. A. & Priess, J. R. APH-1 is a multipass membrane protein essential for the Notch signaling pathway in *Caenorhabditis elegans* embryos. *Proceedings of the National Academy of Sciences of the United States of America* 99, (2002).

Guilmatre, A. et al. Recurrent rearrangements in synaptic and neurodevelopmental genes and shared biologic pathways in schizophrenia, autism, and mental retardation. *Arch Gen Psychiatry* 66, (2009).

Guyant-Marechal, I. et al. Intrafamilial diversity of phenotype associated with app duplication. *Neurology* 71, (2008).

Hardy, J. & Selkoe, D. J. The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. *Science (New York, N.Y.)* 297, (2002).

Hardy, J. A. & Higgins, G. A. Alzheimer's disease: the amyloid cascade hypothesis. *Science (New York, N.Y.)* 256, (1992).

Hardy, J. The Alzheimer family of diseases: many etiologies, one pathogenesis? *Proceedings of the National Academy of Sciences of the United States of America* 94, (1997).

Harold, D. et al. Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. *Nature genetics* 41, (2009).

Heinzen, E. L. et al. Genome-wide scan of copy number variation in late-onset Alzheimer's disease. *Journal of Alzheimer's disease : JAD* 19, (2010).

Hirtz, D. et al. How common are the common neurologic disorders? *Neurology* 68, (2007).

Holtzman, D. M. In vivo effects of ApoE and clusterin on amyloid-beta metabolism and neuropathology. *Journal of molecular neuroscience : MN* 23, (2004).

Hu, Y. et al. Elevated levels of phosphorylated neurofilament proteins in cerebrospinal fluid of Alzheimer disease patients. *Neuroscience letters* 320, (2002).

Iribarren, P., Zhou, Y., Hu, J., Le Yingying & Wang, J. M. Role of formyl peptide receptor-like 1 (FPRL1/FPR2) in mononuclear phagocyte responses in Alzheimer disease. *Immunologic research* 31, (2005).

Itsara, A. et al. Population analysis of large copy number variants and hotspots of human genetic disease. *American journal of human genetics* 84, (2009).

Jack, C. R., Jr et al. Antemortem MRI findings correlate with hippocampal neuropathology in typical aging and dementia. *Neurology* 58, (2002).

JACOBS, P. A. & STRONG, J. A. A case of human intersexuality having a possible XXY sex-determining mechanism. *Nature* 183, (1959).

Jacquemont, S. et al. Mirror extreme BMI phenotypes associated with gene dosage at the chromosome 16p11.2 locus. *Nature* 478, (2011).

Jaeger, L.B. et al. Testing the neurovascular hypothesis of Alzheimer's disease: LRP-1 antisense reduces blood-brain barrier clearance, increases brain levels of amyloid-beta protein, and impairs cognition. *J Alzheimers Dis.* 17, (2009).

Jarrett, J. T., Berger, E. P. & Lansbury, P. T., Jr The carboxy terminus of the beta amyloid protein is critical for the seeding of amyloid formation: implications for the pathogenesis of Alzheimer's disease. *Biochemistry* 32, (1993).

Jellinger, K. A. Traumatic brain injury as a risk factor for Alzheimer's disease. *Journal of neurology, neurosurgery, and psychiatry* 75, (2004).

Jun, G. et al. Meta-analysis confirms CR1, CLU, and PICALM as alzheimer disease risk loci and reveals interactions with APOE genotypes. *Archives of neurology* 67, (2010).

Kallioniemi, A. et al. Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science (New York, N.Y.)* 258, (1992).

Karlsson, R. et al. MAGI1 copy number variation in bipolar affective disorder and schizophrenia. *Biol Psychiatry* 71, (2012).

Karlstrom, H. et al. Variable phenotype of Alzheimer's disease with spastic paraparesis. *Journal of neurochemistry* 104, (2008).

KIDD, M. Paired helical filaments in electron microscopy of Alzheimer's disease. *Nature* 197, (1963).

Kim, J., Basak, J. M. & Holtzman, D. M. The role of apolipoprotein E in Alzheimer's disease. *Neuron* 63, (2009).

Kim, P. M., Korbel, J. O. & Gerstein, M. B. Positive selection at the protein network periphery: evaluation in terms of structural constraints and cellular context. *Proceedings of the National Academy of Sciences of the United States of America* 104, (2007).



Klopocki, E. et al. Complex inheritance pattern resembling autosomal recessive inheritance involving a microdeletion in thrombocytopenia-absent radius syndrome. *Am J Hum Genet* 80, (2007).

Knopman, D. S. et al. Practice parameter: diagnosis of dementia (an evidence-based review). Report of the Quality Standards Subcommittee of the American Academy of Neurology. *Neurology* 56, (2001).

Koolen, A.D. et al. Clinical and molecular delineation of the **17q21.31** microdeletion syndrome. *J Med Genet* 45, (2008).

Kudo, T. et al. Are cerebrovascular factors involved in Alzheimer's disease? *Neurobiology of aging* 21, (2000).

Kumar-Singh, S. et al. Mean age-of-onset of familial alzheimer disease caused by presenilin mutations correlates with both increased Abeta42 and decreased Abeta40. *Human mutation* 27, (2006).

Lambert, J. et al. Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. *Nature genetics* 41, (2009).

Langui, D. et al. Subcellular topography of neuronal Abeta peptide in APPxPS1 transgenic mice. *The American journal of pathology* 165, (2004).

LEJEUNE, J., TURPIN, R. & GAUTIER, M. [Chromosomic diagnosis of mongolism]. *Archives franaises de pdiatrie* 16, (1959).

Levy-Lahad, E. et al. Candidate gene for the chromosome 1 familial Alzheimer's disease locus. *Science (New York, N.Y.)* 269, (1995).

Lin, X. et al. Human aspartic protease memapsin 2 cleaves the beta-secretase site of beta-amyloid precursor protein. *Proceedings of the National Academy of Sciences of the United States of America* 97, (2000).

Lipshutz, R. J., Fodor, S. P., Gingeras, T. R. & Lockhart, D. J. High density synthetic oligonucleotide arrays. *Nature genetics* 21, (1999).

Little, S. P. et al. Zyme, a novel and potentially amyloidogenic enzyme cDNA isolated from Alzheimer's disease brain. *The Journal of biological chemistry* 272, (1997).

Lohmueller, K. E. et al. Proportionally more deleterious genetic variation in European than in African populations. *Nature* 451, (2008).

Lopez, O. L. et al. Reliability of NINCDS-ADRDA clinical criteria for the diagnosis of Alzheimer's disease. *Neurology* 40, (1990).

Luo, W. et al. PEN-2 and APH-1 coordinately regulate proteolytic processing of presenilin 1. *The Journal of biological chemistry* 278, (2003).

Lupski, J. R. Genome structural variation and sporadic disease traits. *Nature genetics*

38, (2006).

Malan, V. & Romana, S. [Diagnosis of chromosomal abnormalities by array CGH in constitutional pathology: the end of the first-line karyotype]. *Arch Pediatr* 19, (2012).

Marlow, L. et al. APH1, PEN2, and Nicastrin increase Abeta levels and gamma-secretase activity. *Biochemical and biophysical research communications* 305, (2003).

Marr, R. A. et al. Neprilysin regulates amyloid Beta peptide levels. *Journal of molecular neuroscience* : MN 22, (2004).

McCarroll, S. A. et al. Integrated detection and population-genetic analysis of SNPs and copy number variation. *Nature genetics* 40, (2008).

McKhann, G. et al. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology* 34, (1984).

McNaughton, D. et al. Duplication of amyloid precursor protein (APP), but not prion protein (PRNP) gene is a significant cause of early onset dementia in a large UK series. *Neurobiology of aging* 33, (2012).

Miller, D.T. et al. Consensus statement: Chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am J Hum Genet* 86, (2010).

Morgan, K. The three new pathways leading to Alzheimer's disease. *Neuropathology and applied neurobiology* 37, (2011).

Motter, R. et al. Reduction of beta-amyloid peptide<sub>42</sub> in the cerebrospinal fluid of patients with Alzheimer's disease. *Annals of neurology* 38, (1995).

Ng, P. C. et al. Genetic variation in an individual human exome. *PLoS genetics* 4, (2008).

NOWELL, P. C. & HUNGERFORD, D. A. Chromosome studies on normal and leukemic human leukocytes. *Journal of the National Cancer Institute* 25, (1960).

Ogawa, K. et al. Localization of a novel type trypsin-like serine protease, neurosin, in brain tissues of Alzheimer's disease and Parkinson's disease. *Psychiatry and clinical neurosciences* 54, (2000).

Perl, D. P. Neuropathology of Alzheimer's disease and related disorders. *Neurologic clinics* 18, (2000).

Pinkel, D. et al. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nature genetics* 20, (1998).

Polvikoski, T. M. et al. Frontal lobe white matter hyperintensities and neurofibrillary pathology in the oldest old. *Neurology* 75, (2010).

Pottier, C. et al. Amyloid- protein precursor gene expression in alzheimer's disease and other conditions. *Journal of Alzheimer's disease : JAD* 28, (2012).

Pottier, C. et al. High frequency of potentially pathogenic SORL1 mutations in autosomal dominant early-onset Alzheimer disease. *Molecular psychiatry* (2012).

Rapoport, S. I. Integrated phylogeny of the primate brain, with special reference to humans and their diseases. *Brain research. Brain research reviews* 15, (1990).

Redon, R. et al. Global variation in copy number in the human genome. *Nature* 444, (2006).

Reitz, C. et al. Meta-analysis of the association between variants in SORL1 and Alzheimer disease. *Archives of neurology* 68, (2011).

Rogaev, E. I. et al. Familial Alzheimer's disease in kindreds with missense mutations in a gene on chromosome 1 related to the Alzheimer's disease type 3 gene. *Nature* 376, (1995).

Rogaeva, E. et al. The neuronal sortilin-related receptor SORL1 is genetically associated with Alzheimer disease. *Nature genetics* 39, (2007).

Rogaeva, E., Kawarai, T. & George-Hyslop, P. S. Genetic complexity of Alzheimer's disease: successes and challenges. *Journal of Alzheimer's disease : JAD* 9, (2006).

Rogers, J. et al. Peripheral clearance of amyloid beta peptide by complement C3-dependent adherence to erythrocytes. *Neurobiology of aging* 27, (2006).

Rovelet-Lecrux, A. et al. A genome-wide study reveals rare CNVs exclusive to extreme phenotypes of Alzheimer disease. *European journal of human genetics : EJHG* 20, (2012).

Rovelet-Lecrux, A. et al. APP locus duplication causes autosomal dominant early-onset Alzheimer disease with cerebral amyloid angiopathy. *Nature genetics* 38, (2006).

Rovelet-Lecrux, A. et al. Deletion of the progranulin gene in patients with frontotemporal lobar degeneration or Parkinson disease. *Neurobiology of disease* 31, (2008).

Rovelet-Lecrux, A. et al. Partial deletion of the MAPT gene: a novel mechanism of FTDP-17. *Human mutation* 30, (2009).

Salminen, A., Ojala, J., Suuronen, T., Kaarniranta, K. & Kauppinen, A. Amyloid-beta oligomers set fire to inflammasomes and induce Alzheimer's pathology. *Journal of cellular and molecular medicine* 12, (2008).

Saugier-Verber, P. et al. Detection of heterozygous SMN1 deletions in SMA families using a simple fluorescent multiplex PCR method. *Journal of medical genetics* 38, (2001).

Schellenberg, G. D. et al. Genetic linkage evidence for a familial Alzheimer's disease

locus on chromosome 14. *Science* (New York, N.Y.) 258, (1992).

Schmechel, D. E. et al. Increased amyloid beta-peptide deposition in cerebral cortex as a consequence of apolipoprotein E genotype in late-onset Alzheimer disease. *Proceedings of the National Academy of Sciences of the United States of America* 90, (1993).

Schouten, J. P. et al. Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic acids research* 30, (2002).

Sebat, J. et al. Large-scale copy number polymorphism in the human genome. *Science* (New York, N.Y.) 305, (2004).

Selkoe, D. J. Amyloid beta-protein precursor: new clues to the genesis of Alzheimer's disease. *Current opinion in neurobiology* 4, (1994).

Shah, S. et al. Nicastrin functions as a gamma-secretase-substrate receptor. *Cell* 122, (2005).

Shankar, G. M. et al. Natural oligomers of the Alzheimer amyloid-beta protein induce reversible synapse loss by modulating an NMDA-type glutamate receptor-dependent signaling pathway. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 27, (2007).

Sharp, A.J. et al. A recurrent 15q13.3 microdeletion syndrome associated with mental retardation and seizures. *Nat Genet* 40, (2008).

Sherrington, R. et al. Alzheimer's disease associated with mutations in presenilin 2 is rare and variably penetrant. *Human molecular genetics* 5, (1996).

Sherrington, R. et al. Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. *Nature* 375, (1995).

Shiraishi, H. et al. PEN-2 enhances gamma-cleavage after presenilin heterodimer formation. *Journal of neurochemistry* 90, (2004).

Slegers, K. et al. APP duplication is sufficient to cause early onset Alzheimer's dementia with cerebral amyloid angiopathy. *Brain : a journal of neurology* 129, (2006).

Steiner, H. et al. A pathogenic presenilin-1 deletion causes aberrant Abeta 42 production in the absence of congophilic amyloid plaques. *The Journal of biological chemistry* 276, (2001).

Sudmant, P. et al. Diversity of Human Copy Number Variation and Multicopy Genes. *Science* 330, (2010).

Swaminathan, S. et al. Genomic Copy Number Analysis in Alzheimer's Disease and Mild Cognitive Impairment: An ADNI Study. *International journal of Alzheimer's disease* 2011, 729478 (2011).

Tabira, T., de Chui, H., Nakayama, H., Kuroda, S. & Shibuya, M. Alzheimer's disease with spastic paresis and cotton wool type plaques. *Journal of neuroscience research* 70, (2002).

Tanzi, R. E. et al. Amyloid beta protein gene: cDNA, mRNA distribution, and genetic linkage near the Alzheimer locus. *Science (New York, N.Y.)* 235, (1987).

TERRY, R. D. THE FINE STRUCTURE OF NEUROFIBRILLARY TANGLES IN ALZHEIMER'S DISEASE. *Journal of neuropathology and experimental neurology* 22, (1963).

Thornton, E., Vink, R., Blumbergs, P. C. & van den Heuvel, C. Soluble amyloid precursor protein alpha reduces neuronal injury and improves functional outcome following diffuse traumatic brain injury in rats. *Brain research* 1094, (2006).

Tiffany, H. L. et al. Amyloid-beta induces chemotaxis and oxidant stress by acting at formylpeptide receptor 2, a G protein-coupled receptor expressed in phagocytes and brain. *The Journal of biological chemistry* 276, (2001).

Tolia, A. & de Strooper, B. Structure and function of gamma-secretase. *Seminars in cell developmental biology* 20, (2009).

Uéda, K. et al. Molecular cloning of cDNA encoding an unrecognized component of amyloid in Alzheimer disease. *Proceedings of the National Academy of Sciences of the United States of America* 90, (1993).

Vallet, P. G. et al. A comparative study of histological and immunohistochemical methods for neurofibrillary tangles and senile plaques in Alzheimer's disease. *Acta neuropathologica* 83, (1992).

Vandermeeren, M. et al. Detection of tau proteins in normal and Alzheimer's disease cerebrospinal fluid with a sensitive sandwich enzyme-linked immunosorbent assay. *Journal of neurochemistry* 61, (1993).

Vassar, R. et al. Beta-secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE. *Science (New York, N.Y.)* 286, (1999).

Vepsäläinen, S. et al. Increased expression of Abeta degrading enzyme IDE in the cortex of transgenic mice with Alzheimer's disease-like neuropathology. *Neuroscience letters* 438, (2008).

Vepsäläinen, S. et al. Insulin-degrading enzyme is genetically associated with Alzheimer's disease in the Finnish population. *Journal of medical genetics* 44, (2007).

Vetrivel, K. S. & Thinakaran, G. Amyloidogenic processing of beta-amyloid precursor protein in intracellular compartments. *Neurology* 66, (2006).

Visser, P. J., Scheltens, P., Pelgrim, E. & Verhey, F. R. Medial temporal lobe atrophy and APOE genotype do not predict cognitive improvement upon treatment with rivastigmine in Alzheimer's disease patients. *Dementia and geriatric cognitive disorders* 19, (2005).

Wallon, D. et al. The French Series of Autosomal Dominant Early Onset Alzheimer's Disease Cases: Mutation Spectrum and Cerebrospinal Fluid Biomarkers. *Journal of Alzheimer's disease : JAD* (2012).

Walters, R.G. et al. A new highly penetrant form of obesity due to deletions on chromosome 16p11.2. *Nature* 463, (2010).

Whitehouse, P. J., Price, D. L., Clark, A. W., Coyle, J. T. & DeLong, M. R. Alzheimer disease: evidence for selective loss of cholinergic neurons in the nucleus basalis. *Annals of neurology* 10, (1981).

Wolfe, M. S. When loss is gain: reduced presenilin proteolytic function leads to increased Abeta42/Abeta40. *Talking Point on the role of presenilin mutations in Alzheimer disease. EMBO reports* 8, (2007).

Wu, Z. et al. Role of the MEOX2 homeobox gene in neurovascular dysfunction in Alzheimer disease. *Nature medicine* 11, (2005).

Yan, R. et al. Membrane-anchored aspartyl protease with Alzheimer's disease beta-secretase activity. *Nature* 402, (1999).

Ying, G. et al. Humanin, a newly identified neuroprotective factor, uses the G protein-coupled formylpeptide receptor-like-1 as a functional receptor. *Journal of immunology* 172, (2004).

Yu, G. et al. Nicastrin modulates presenilin-mediated notch/glp-1 signal transduction and betaAPP processing. *Nature* 407, (2000).

Zekanowski, C. et al. Two novel presenilin 1 gene mutations connected with frontotemporal dementia-like clinical phenotype: genetic and bioinformatic assessment. *Experimental neurology* 200, (2006).

Zeng, F. et al. Array-MLPA: comprehensive detection of deletions and duplications and its application to DMD patients. *Human mutation* 29, (2008).

Zhang, F. et al. Copy number variation in human health, disease, and evolution. *Annu Rev Genomics Hum Genet* 10, (2009).