



Genetic analysis of familial Alzheimer's disease, primary lateral sclerosis and paroxysmal kinesigenic dyskinesia: a tool to uncover common mechanistic points

Author
Jacek Szymański

Doctoral programme in neurosciences

Director
Jordi Pérez-Tur

Tutor
Alicia Salvador
Fernández-Montejo

November 2020

Dr Jordi Pérez-Tur, Investigador Científico del Instituto de Biomedicina de Valencia (IBV-CSIC), en calidad de Director de la tesis doctoral de Jacek Szymański, adscrito al Programa de Doctorado en Neurociencias de la Universitat de València.

CERTIFICA

Que la tesis titulada “Genetic analysis of familial Alzheimer’s disease, primary lateral sclerosis and paroxysmal kinesigenic dyskinesia: a tool to uncover common mechanistic points” se ha desarrollado bajo su dirección y supervisión, y que el trabajo de investigación realizado y la memoria del mismo, ha sido elaborada por el doctorado y cumple los requisitos científicos y formales para proceder al acto de defensa de la Tesis Doctoral. Y para que conste, en el cumplimiento de la legislación presente, firman el presente certificado en València, 2020.

Dr Jordi Pérez-Tur
Director

Dra Alicia Salvador
Fernández-Montejo
Tutor Académico

Jacek Szymański
Doctorando

ACKNOWLEDGEMENTS

This thesis is dedicated to Hanna Szymańska, the person to whom I owe my decision to pursue a doctoral degree and whose knowledge of science and fullest support I could always count on.

I would like to express my deepest appreciation and gratitude to my supervisor and director of this thesis, Dr Jordi Pérez-Tur, for his valuable advice, unparalleled support, constructive criticism and guidance. He accepted me under his wing, created a wonderful work environment and is undoubtedly the best group leader I have encountered in my scientific career.

I am deeply indebted to Fernando Cardona, without whose support and nurturing, this thesis would not have been possible. Always ready with helpful advice and practical suggestions, I am extremely grateful to him, and I am fortunate to have met him.

I am thankful to CIBERNED (Centro de Investigación Biomédica en Red de Enfermedades Neurodegenerativas) for funding this research into such important neurodegenerative diseases. I am also grateful to all my colleagues at IBV-CSIC (Instituto de Biomedicina de Valencia - Consejo Superior de Investigaciones Científicas) who supported me with great advice and camaraderie through the years. Especially Jessica Valdivia, Conchi Rubio and Guillermo Casero, who helped me with my work, brightening up my day with great humour.

I would like to thank Małgorzata Wiweger and Jacek Kuźnicki who accepted me for a training stay at the IIMCB (International Institute of Molecular and Cell Biology) in Warsaw, which greatly improved my knowledge and abilities. I would like to acknowledge the assistance and collaboration of Ana Espinosa and Teresa Caballero Vizcaíno from the IBMCP-CSIC (Institute for Plant Molecular and Cell Biology), and Jerónimo Bravo and Manuel Soriano from IBV-CSIC. Also the contribution of clinicians: Juan Francisco Vázquez Costa and Juan A. Burguera from Servicio de Neurología, Hospital Universitario y Politécnico de La Fe, and Àngel Pérez Sempere from Hospital General Universitario de Alicante.

Special thanks to my family, Maciej, Hanna, Katarzyna, Michał, Tomasz, Keri, Dickson, Mały and Karo for unwavering support and endless good cheer and generosity. In particular, I want to thank Agnieszka, whose support cannot be overestimated, who kept me motivated, patient, and most of all happy. Finally, I am grateful to my friends, like Luke, for a healthy work–life balance.

Wiara w daktyla nie jest daktylem samym w sobie

– Hanna Szymańska

SUMMARY

Introduction

Neurological diseases can have a wide spectrum of phenotypes, causing disruptions in daily life activities and, often, a progressive disease ending in loss of life. Molecular mechanisms underlying many of these diseases can have common pathways and depending on the gene in question, its dysfunction may lead to a variety of conditions.

Alzheimer's disease (AD) is the most common neurodegenerative dementia in the World with an estimate of 50 million people living with dementia worldwide. Symptoms of AD include difficulty remembering names or recent events, apathy, depression, disorientation, confusion, difficulty speaking, walking and swallowing. The canonical division of AD into early onset (EOAD) and late onset (LOAD) is set at 65 years of age. Both are characterised by intracellular hyperphosphorylated tau protein aggregates called neurofibrillary tangles (NFTs) and extracellular senile plaques composed primarily of clumps of amyloid- β ($A\beta$) peptide. The AD pathology starts at the entorhinal cortex, a region of the medial temporal lobe and causes neuron death leading to atrophy of brain tissue. AD is a multifactorial disease with a dichotomous pattern of inheritance with approximately 70 % of the causes being genetic and the rest environmental. Mutations in genes: *APP*, *PSEN1* and *PSEN2*, are the most common causes of EOAD and allele $\epsilon 4$ of *APOE* is a well-established risk factor of LOAD. Three main pathways encapsulate most of the AD genetic risk factors: the immune response, lipid metabolism and endocytosis. Among the genes involved in immune response, *CR1* has been associated with AD through genome-wide association studies (GWAS) and is an important part of the complement system, and significant for this work. A particular CR1 isoform, CR1*2, is expressed at lower levels than isoform CR1*1 and it is speculated it affects lower $A\beta$ clearance and dysregulation of the complement system.

Primary lateral sclerosis (PLS) is considered as part of the amyotrophic lateral sclerosis (ALS) pathological spectrum. ALS, although considered a rare disease, is the most common motor neuron disease (MND). PLS is characterised by spinobulbar spasticity caused by upper motor neurons (UMNs) degeneration, while ALS is characterized by degeneration of both UMNs and lower motor neurons (LMNs) at spinal and bulbar level, causing limb paralysis, dysarthria, dysphagia and fatal respiratory failure. Clinical features of PLS include spasticity, slight weakness in the lower limbs, adult-onset, progressive

course, duration of longer than 4 years and pseudobulbar symptoms. Behavioural and cognitive deficits may occur with ALS, ranging from mild, moderate to frontotemporal dementia (FTD) which was also seen in patients with PLS. The genetic basis of PLS is not well understood although mutations in genes *SPG7* and *TBK1* were reported in patients affected by familial PLS. Mutations in *SOD1* gene, encoding an antioxidant enzyme, and in the protein encoded by *C9orf72* gene are the most common causes of ALS. ALS pathogenesis may be caused by glutamate excitotoxicity, mitochondrial dysfunction, impaired structure or transport in axons and oxidative stress.

Paroxysmal kinesigenic dyskinesia (PKD), the most common paroxysmal movement disorder, is characterized by a range of involuntary movements triggered by sudden motion. With onset in childhood or adolescence its clinical features include recurrent attacks involving chorea, athetosis, dystonic postures or ballismus. Severity of these attacks typically decreases with age. Mutations in gene *PRRT2* are thus far the only cause of this disorder. *PRRT2* is known to interact with proteins SNAP-25, SYT1 and SYT2 in the presynaptic membrane of neurons, which are involved in signalling in nerve cells. The lack of *PRRT2* often caused through nonsense-mediated mRNA decay (NMD) pathway, due to premature termination codons (PTCs) in the transcript, is said to be the common molecular mechanism involved in haploinsufficiency causing PKD.

Objectives

The main objective of this work is to comprehend the etiopathogenesis of three neurological diseases affecting three distinct Spanish families and to study the genetic and molecular mechanisms affecting these diseases.

Materials and methods

DNA samples were collected from members of three Spanish families: UGM037 (11 individuals) affected by AD, UGM471 (7 individuals) affected by PLS and UGM478 (7 individuals) affected by PKD. A control population of healthy individuals and individuals with AD was also available and all the samples were collected with approval of the corresponding institutional review boards of the corresponding hospitals with signed informed consent from patients. The bacterial strain used for all necessary manipulations was *Escherichia coli DH5 α* and the human cell line was SH-SY5Y.

Whole-exome sequencing (WES) was used as the tool to identify variants within the exome of the patients for further investigation. Through base-calling and image

analysis, read alignment and SNP calling the data could be transformed into a workable database of variants. Filtering and prioritization followed by Sanger sequencing validation of the results was used to identify the most interesting variants, possibly involved in causing the disease. Public databases such as Collaborative Spanish Variant Server (CSVS) or Genome Aggregation Database (gnomAD) were used for variant assessment. The variant frequency was also verified through allele-specific PCR (ASPCR) performed on the control populations.

After identifying the most interesting variants, plasmid manipulations were performed to obtain specific cDNAs, corresponding to sequences of the genes with these variants, in specific expression vectors. Site-directed mutagenesis was used for introduction of single nucleotide variants (SNV), FLAG epitope and enzyme restriction sites. Subcloning was used to transfer specific cDNAs to corresponding expression vectors. *In silico* analysis was performed using HOPE, I-Mutant 2.0 and ConSurf web services. That way the influence of a specific variant on protein function and stability, as well as whether a specific residue is conserved, could be established.

In the case of genes involved in AD, genotyping of *APOE* and *CRI* was important and performed using PCRs. In the case of *APOE*, the PCR was followed by enzyme restriction and analysis of band distribution on an agarose gel. In the case of *CRI*, the PCR was followed by Sanger sequencing or analysis of presence or absence of bands on an agarose gel.

SH-SY5Y cells were maintained in a complete growth medium and Lipofectamine 2000 Transfection Reagent was used for transfecting them with plasmids corresponding to specific experiments. One experiment was to compare the mRNA and protein levels of specific genes affected by the identified variants. In another, the same comparison was made after inhibition of the NMD pathway, using NMDI14. In order to measure mRNA levels, the mRNA was extracted from transfected SH-SY5Y cells after approximately 48h of incubation and by reverse transcription cDNAs were obtained. These were used in quantitative PCR (qPCR) with specifically designed primers to amplify the reverse transcribed transcript with appropriate controls. In order to measure protein levels, the proteins were extracted from transfected SH-SY5Y cells after approximately 48h of incubation and used for Western Blot (WB) analysis. Housekeeping gene actin was used as control.

For functional analysis of ADPRH protein and its variants, the proteins were extracted from transfected SH-SY5Y cells and subjected to purification using GST-tagged

protein affinity column. After sample dialysis and concentration the proteins could be used for an activity assay. This involved cholera toxin (CT) ADP-ribosylation of substrate and further ADP-ribose cleavage by ADPRH and its variants. The samples were then loaded onto an ultraperformance liquid chromatography (UPLC) column which allowed for separation, identification, and quantification of components based on data from a detector measuring absorbance at 260 nm. Also co-immunoprecipitation was used with protein extracts from SH-SY5Y cells co-transfected with ADPRH and ADPRHL1 in order to ascertain whether protein-protein interaction takes place.

Statistical analysis was performed using the student's t test and one-way or two-way ANOVA to test differences between group means, with a *post hoc* Tukey multiple comparisons of means test. R software was used to perform the tests.

Results and discussion

The filtration, prioritization and Sanger sequencing validation identified variant rs764542666 in gene *CR1* encoding a PTC c.C406T p.R136* ($CR1^{R136*}$) as the likely cause of AD in family UGM037. The clinical data suggested a LOAD in the family, which would be more in line with risk factors rather than causative mutations. The family members did possess the *APOE* $\epsilon 4$ allele, however remarkably, genotyping revealed a healthy member of the family with a $\epsilon 4/\epsilon 4$ *APOE* genotype at age 88. The $CR1^{R136*}$ variant segregated with the disease in the pedigree, not affecting any of the non-AD members. As $CR1^{*2}$ isoform of *CR1*, which is expressed at lower levels than $CR1^{*1}$, was previously associated with higher risk of LOAD, a possible NMD of $CR1^{R136*}$ transcript could cause haploinsufficiency and a similar effect. Study of mRNA and protein levels revealed the $CR1^{R136*}$ to be expressed at lower levels than the wild-type ($CR1^{WT}$). These levels would then significantly increase after treatment with NMD inhibitor, suggesting involvement of this pathway in $CR1^{R136*}$ transcript destruction. Genotyping of $CR1^{*2}$ isoform and rs3818361 in *CR1* in samples from UGM037 family, showed that none of them had $CR1^{*2}$ isoform and rs3818361 did not segregate with the disease, not being present in any of the samples from patients. Both of these were genotyped as they were previously described as associated with AD. $CR1^{*2}$ isoform, as mentioned before, is expressed at lower levels than $CR1^{*1}$ thus possibly affecting lower $A\beta$ clearance and dysregulation of the complement system. The rs3818361 was genotyped to verify whether it may be the real culprit behind the disease, being in linkage disequilibrium with $CR1^{R136*}$. Children of the patients and the one healthy family member, were not diagnosed with AD, however their DNA samples

were collected at ages between 50 and 57. They may be at risk of developing the disease at a later age, especially as all except for one of them had one copy of the *APOE* ϵ 4 allele. Also three out of seven of them had the $CR1^{R136^*}$ variant. Although *in vitro* overexpression of a protein in a cellular model has its limitations, the molecular mechanism behind AD in family UGM037 seems to be haploinsufficiency caused by NMD pathway destruction of $CR1^{R136^*}$ transcript. Further functional studies on the effects of $CR1^{R136^*}$ would be recommended and a future follow up with the younger members of the family. To the best of my knowledge, $CR1^{R136^*}$ (rs764542666) would be the first known AD causative mutation in gene *CR1*.

Available clinical data supported the diagnosis of PLS in UGM471 family members. Through WES data filtration, prioritization and Sanger sequencing validation two variants were identified which could possibly cause the symptoms in the family. One was a novel variant c.G884C p.R295P in *ADPRH* ($ADPRH^{R295P}$) and the other a previously described mutation c.T497C p.L166P in *PSEN1* (rs63750265; $PSEN1^{L166P}$). The latter was discovered within the UGM471 family genome recently, therefore they will be discussed chronologically, from discovery.

ADPRH is a ubiquitous protein found in cytoplasm of both mice and humans. *ADPRH* hydrolyses the *N*-glycosidic bond between arginine and ADP-ribose, cleaving ADP-ribose from substrate in the ADP-ribosylation cycle. The reaction occurs to be specific to mono-ADP-ribosylated substrates, due to the protein's structure. ADP-ribosylation cycle is very important for cell regulation. The $ADPRH^{R295P}$ variant was predicted to be deleterious by SIFT and Polyphen. It was not found in any public database and Sanger sequencing confirmed its segregation with the disease in UGM471 family pedigree. However, two healthy individuals from this family also carried $ADPRH^{R295P}$. ASPCR frequency assessment of $ADPRH^{R295P}$ in control population showed none of the 196 non-PLS individuals were carriers. The arginine in position 295 in *ADPRH* was found to be a conserved residue and a change to a proline could affect the function or stability of the protein significantly. Reassessment of WES data showed a paralog of *ADPRH*, *ADPRHL1* to be affected by an $ADPRHL1^{L294R}$ variant, also predicted to be deleterious. This variant segregated with the disease in the family, not being present in the healthy individuals. *ADPRH* together with *ADPRHL1* are speculated to be involved in actin filaments assembly and modulation of actin polymerisation may be involved in disruption of the nucleocytoplasmic transport which is important in ALS pathogenesis. Thus the $ADPRH^{R295P}$ variant may have a low penetrance, causing the disease only in certain

members of the family, or may be accompanied by ADPRHL1^{L294R} variant, to affect the patients. Co-immunoprecipitation did not show an interaction between the two proteins, however they may work in unison without a direct interaction. Activity assay for ADPRH^{R295P}, where cholera toxin (CT) was used to ADP-ribosylate a substrate and wild-type ADPRH (ADPRH^{WT}) and its variants (ADPRH^{R295P} and ADPRH^{D55A/D56A}) were used to cleave ADP-ribose, showed that ADPRH^{R295P} had a similar activity efficiency as ADPRH^{WT}. The ADPRH^{D55A/D56A} variant where the essential active site residues were changed, did not show hydrolase activity. Although this assay showed no diminishing of the activity of variant ADPRH^{R295P}, it did not take into account the possible destabilising effect of the variant. The RT-qPCR assay to measure mRNA levels of this variant showed that they did not vary from the wild-type, however, the protein levels measured with WB proved to be significantly lower than the ones for ADPRH^{WT} or for other variants examined (ADPRH^{R295Q}, ADPRH^{D55A/D56A}). Only very rare truncating variant ADPRH^{R295*} which has been previously described, proved to have very low mRNA levels and no protein could be detected. This suggest it may be affected by the NMD pathway. While it is difficult to say whether variant ADPRH^{R295P} affects the disease in family UGM471 it is important to know its strong effect on the protein stability for the further study of little studied mono-ADP-ribosylation.

Mutation c.T497C p.L166P in *PSEN1* (rs63750265) was recently discovered in family UGM471. Although mutations in *PSEN1* are commonly causing EOAD, there have been reports associating *PSEN1* with PLS and ALS. PSEN1^{A431E} and PSEN1^{L381V} are two examples of mutations associated with PLS and AD, causing an early onset at around 40 years of age. Mutation PSEN1^{L166P} is known to cause AD and spastic paraparesis at an early age, with the first case found being a 15 year old girl. Other mutations at position L166 in *PSEN1* were also found to be associated with AD (PSEN1^{L166V}, PSEN1^{L166R} (rs63750265), PSEN1^{L166H} (rs63750265), and PSEN1^{L166del} (rs63751458)). Most of them were associated with cognitive impairment, although some were associated with motor symptoms. PSEN1^{L166P} mutation causes partial loss of γ -secretase cleavage function and increases the A β ₄₂/A β ₄₀ ratio by reducing the A β ₄₀ levels. Also PSEN1 functions as endoplasmic reticulum (ER) Ca²⁺ leak channels, and the PSEN1^{L166P} mutation disrupts that function. Recent discovery of the PSEN1^{L166P} mutation in the UGM471 family did not allow for a more thorough investigation of its effects, but it was confirmed to segregate with the disease in the pedigree. The aggressive nature of this mutation, the early age of onset and the motor symptoms, strongly suggest it is PSEN1^{L166P} which causes the disease in the family. While

it is not associated with ALS nor PLS, the patients may have been also affected by other environmental or genetic factors to produce this phenotype. At the same time reports of PLS patients with mutations in *PSENI* exist. Either the effect of the *PSENI* mutations is heterogeneous enough to cause these different disorders, the disorders are much more related due to the molecular mechanisms that cause them, or the effect of *PSENI* mutation is affected by interaction with other proteins. Certainly, common neurodegeneration-linked genes should be looked at when identifying a possible cause of a disease, regardless of which disease they are associated with. Patients with PLS, ALS or spastic paraparesis should be investigated for *PSENI* mutations.

In PKD, the filtration, prioritization and Sanger sequencing validation identified a novel variant c.C316T p.Q106* in *PRRT2* ($PRRT2^{Q106*}$) as the likely cause of the disease in family UGM478. Clinical data for the family support the diagnosis of PKD. The assessment of the WES results lead to a strong suspicion of $PRRT2^{Q106*}$ as being the causative mutation of the disease. Mutations in *PRRT2* are the only known cause of PKD thus far, and often they are truncating mutations leading to haploinsufficiency due to NMD pathway destruction of PTC carrying transcript. $PRRT2^{Q106*}$ was not found in any database therefore the only frequency data was obtained through our ASPCR assay for a control population of 192 samples. None of them carried this variant, yet it segregated with the disease in the pedigree, not being present in any of the non-PKD family members. In accordance with previous results, significantly lower mRNA levels for the variants with PTCs ($PRRT2^{Q106*}$, $PRRT2^{Q163*}$, $PRRT2^{Q250*}$), in comparison with the wild-type ($PRRT2^{WT}$), were found in a cellular model with *PRRT2* and its variants overexpressed. $PRRT2^{Q163*}$ and $PRRT2^{Q250*}$ were chosen as controls as previously described variants in *PRRT2*. Inhibition of the NMD pathway by treatment of transfected SH-SY5Y cells with NMDI14, increased the mRNA levels for $PRRT2^{Q106*}$ and $PRRT2^{Q163*}$ and showed an increasing trend for $PRRT2^{Q250*}$. This suggests that the NMD pathway may in fact be the culprit behind mRNA decay prompted by PTCs in the variants studied. Interestingly protein levels of the novel variant $PRRT2^{Q106*}$ were undetectable with WB before and after NMDI14 treatment, while the other variants studied ($PRRT2^{Q163*}$ and $PRRT2^{Q250*}$) had significantly lower protein levels than $PRRT2^{WT}$ before treatment and increased levels after. This may be due to the proline-rich regions being important for protein stability, as $PRRT2^{Q106*}$ has a PTC before that region and $PRRT2^{Q163*}$ and $PRRT2^{Q250*}$ in the middle and after it. These results suggest the novel variant $PRRT2^{Q106*}$ is probably the cause of PKD in the UGM478 Spanish family. The molecular mechanisms responsible for the

affliction may be the NMD pathway causing decay of the transcript leading to haploinsufficiency. Lack of *PRRT2* in turn causes hyperexcitability through dysregulated neurotransmitter release and hyperactivity of Na⁺ channels.

Common or related pathological molecular mechanisms may affect neurological disorders, traditionally considered as unrelated, in the intricate network of the nervous system. In this work I have outlined some of such putative mechanisms. Single nucleotide variants which may affect different phenotypes through partial loss-of-function due to protein destabilisation or haploinsufficiency due to NMD.

The estimated possible number of human haploinsufficient genes is 12,443 out of approximately 22,000. While the total number of human genes is a matter of debate and further study, their estimation indicates that we can expect a great number of genes where protein level dose effect may be essential. Haploinsufficiency is important in neurological disorders. Recently *C9orf72* was found to be haploinsufficient in ALS/FTD due to the GGGGCC repeat expansion.

In this work, I am postulating that *CR1* and *PRRT2* are haploinsufficient in Spanish families with AD and PKD respectively. While it is established that most mutations in *PRRT2* lead to loss-of-function and haploinsufficiency, to my knowledge there is no such reports on *CR1*. Haploinsufficiency, therefore, emerges as a common factor between these and other neurological diseases. Furthermore, the molecular mechanism behind the *CR1* and *PRRT2*-related haploinsufficiency seems to be the NMD elicited by SNVs encoding PTCs, demonstrating a common molecular mechanism in distinct neurological diseases.

Loss-of-function is strictly related to haploinsufficiency which is a dominant phenotype in organisms heterozygous for such alleles. Although variant *PSEN1*^{L166P} (rs63750265) was not found to cause haploinsufficiency, it affects a partial loss of γ -secretase cleavage and ER Ca²⁺ leak channel function. The variant *ADPRH*^{R295P}, identified in the same family, is shown to significantly destabilize the protein, drastically affecting its levels. This in turn may impede its function. Whether *ADPRH*^{R295P} variant in heterozygosis is in fact deleterious remains to be seen, depending on its tolerance to decreased protein dose. However, loss-of-function, whether full or partial, encompasses the underlying molecular mechanisms of the SNVs described in this work, which contribute to independent neurological diseases.

Discussing the molecular mechanisms in AD and PLS in the two Spanish families, and the involvement of *CR1*^{R136*} (rs764542666) and *PSEN1*^{L166P} (rs63750265) variants in disease pathogenesis, it is important not to omit other possibly contributing factors. While

CR1^{R136*} may be a causative mutation in the family, its members had a high incidence of *APOE* ε4. In families with *APP* mutations, the incidence of *APOE* ε4 was related to an earlier age of onset, while the incidence of *APOE* ε2, with a later age of onset, with regards to *APOE* ε3. Interestingly, PSEN1^{E318G} variant is related to an increased risk of AD, dependent on *APOE* ε4. While otherwise PSEN1^{E318G} was considered non-pathogenic, its interaction with *APOE* ε4 increased Aβ deposition, causing a faster cognitive decline and neurodegeneration. Thus, while carrying CR1^{R136*} variant may be sufficient to develop AD, it is also probable that the members of the Spanish family studied, were affected solely by the *APOE* ε4 risk factor or a combination of the two.

Similarly, other factors, whether environmental or genetic, may affect symptoms developed by the family with PLS. Although PSEN1^{L166P} seems to be responsible for the phenotype experienced by the patients, their symptoms differ from the more canonical AD features related to this variant. This divergent disease expression may be due to *PSEN1* gene pleiotropy, however, it may also be due to other contributing factors. Environmental factors have been found to play an important role in ALS and they cannot be disregarded in a familial disease. Here, I propose a genetic factor which may contribute to the dissimilar symptoms experienced by the members of this family. Novel variant ADPRH^{R295P} may have an effect on the disease development, destabilising actin filaments in the presence of ADPRHL1^{L294R} variant, prompting a phenotype closer to PLS, together with the aggressive PSEN1^{L166P} mutation. The complexity of neurological diseases comes, in part, from a cumulative nature of defects that cause them, and thus it is always essential to search for other factor which may add to the observed phenotype. Further studies into the effects ADPRH variants may have on neurological diseases are needed as it contributes to the still poorly understood, but very important mono-ADP-ribosylation.

The general limitations of the *in vitro* overexpression of a protein in a cellular model apply in the entire study. Differences between the cellular model and the corresponding cells in the organism, problems with establishing appropriate microenvironment, such as interactions with other cells, or the fact that the protein is artificially overexpressed in naturally unavailable amounts. Further functional studies may be needed for all the described variants.

To conclude, variant rs764542666 in gene *CR1* encoding a PTC c.C406T p.R136* is the likely cause of AD in a Spanish family UGM037, based on WES and genetic expression study. NMD pathway provoked haploinsufficiency of CR1 is the probable molecular mechanism behind the disease. Variant rs764542666 is probably the first known

AD causative mutation in *CRI*, encouraging research into the rare truncating variants in this gene.

Mutant rs63750265 in gene *PSENI* encoding a missense mutation c.T497C p.L166P is the likely cause of PLS in a Spanish family UGM471, based on WES study, segregation analysis and previous knowledge, raising questions on pleiotropic effects of the mutation. The molecular mechanisms behind mutant rs63750265 causing the disease in family UGM471 are probably loss of γ -secretase cleavage function, increase of A β ₄₂/A β ₄₀ ratio and impairment of ER Ca²⁺ leak channel function. Novel variant in gene *ADPRH* encoding a missense variant c.G884C p.R295P strongly destabilizes the protein while not affecting its function, shedding light on the study of mono-ADP-ribosylation.

Novel variant in gene *PRRT2* encoding a PTC c.C316T p.Q106* is the likely cause of PKD in a Spanish family UGM037, based on WES and genetic expression study. This work supports the hypothesis of NMD pathway provoking haploinsufficiency of PRRT2 as the molecular mechanism behind PKD.

RESUMEN

Introducción

Las enfermedades neurológicas pueden tener un amplio espectro de fenotipos, causando interrupciones en las actividades de la vida diaria y, a menudo, una enfermedad progresiva que termina en la pérdida de la vida. Los mecanismos moleculares que subyacen a muchas de estas enfermedades pueden tener vías comunes y, dependiendo del gen en cuestión, su disfunción puede llevar a una variedad de condiciones.

La enfermedad de Alzheimer (EA) es la demencia neurodegenerativa más común en el mundo, con una estimación de 50 millones de personas que viven con demencia en todo el mundo. Los síntomas de la EA incluyen dificultad para recordar nombres o eventos recientes, apatía, depresión, desorientación, confusión, dificultad de hablar, caminar y tragar. La división canónica de la EA en inicio temprano (EOAD por sus siglas en inglés) e inicio tardío (LOAD por sus siglas en inglés) se establece en los 65 años de edad. Ambos se caracterizan por agregados de proteína tau hiperfosforilada intracelular llamados ovillos neurofibrilares (NFT por sus siglas en inglés) y placas seniles extracelulares compuestas principalmente por grupos de péptido amiloide- β ($A\beta$). La patología de EA comienza en la corteza entorrinal, una región del lóbulo temporal medial, y causa la muerte de las neuronas provocando la atrofia del tejido cerebral. La EA es una enfermedad multifactorial con un patrón de herencia dicotómico. Aproximadamente el 70 % de las causas son genéticas y el resto ambientales. Mutaciones en los genes: *APP*, *PSEN1* y *PSEN2*, son las causas más comunes de EOAD y el alelo $\epsilon 4$ de *APOE* es un factor de riesgo bien establecido de LOAD. Tres vías principales contienen la mayoría de los factores de riesgo genéticos de EA: la respuesta inmune, el metabolismo lipídico y la endocitosis. Entre los genes involucrados en la respuesta inmune, *CR1* se ha asociado con la EA mediante estudios de asociación de todo el genoma (GWAS por sus siglas en inglés). *CR1* forma un parte importante del sistema del complemento y es significativo para este trabajo. Una isoforma de *CR1*, *CR1*2*, se expresa en niveles más bajos que la isoforma *CR1*1* y se especula que provoca un menor aclaramiento de $A\beta$ y una desregulación del sistema del complemento.

La esclerosis lateral primaria (ELP) se considera una parte del espectro patológico de la esclerosis lateral amiotrófica (ELA). La ELA, aunque se considera una enfermedad rara, es la enfermedad de la neurona motora (ENM) más común. La ELP se caracteriza por

la espasticidad espinobulbar causada por la degeneración de las neuronas motoras superiores (NMS), mientras que la ELA se caracteriza por la degeneración tanto de las NMS como de las neuronas motoras inferiores (NMI) a nivel espinal y bulbar, lo que causa parálisis de las extremidades, disartria, disfagia y muerte por insuficiencia respiratoria. Características clínicas de ELP incluyen la espasticidad, ligera debilidad en las extremidades inferiores, inicio en adultos, curso progresivo, duración de más de 4 años y los síntomas pseudobulbares. Los déficits cognitivos y conductuales pueden ocurrir con ELA, que van desde la demencia leve, moderada a la demencia frontotemporal (DFT) que también se observó en pacientes con ELP. La base genética de la ELP no se conoce bien, aunque se conocen mutaciones en los genes *SPG7* y *TBK1* en pacientes afectados por ELP familiar. Las mutaciones en el gen *SOD1*, que codifica una enzima antioxidante, y en la proteína codificada por el gen *C9orf72* son las causas más comunes de ELA. La patogenia de la ELA puede ser causada por excitotoxicidad del glutamato, disfunción mitocondrial, alteración de la estructura o transporte en los axones y estrés oxidativo.

La discinesia paroxística cinesigética (PKD por sus siglas en inglés), el trastorno del movimiento paroxístico más común, se caracteriza por una variedad de movimientos involuntarios desencadenados por movimientos repentinos. Con inicio en la niñez o adolescencia, sus características clínicas incluyen ataques recurrentes que involucran corea, atetosis, posturas distónicas o balismo. La gravedad de estos ataques suele disminuir con la edad. Las mutaciones en el gen *PRRT2* son, hasta ahora, la única causa conocida de este trastorno. Se sabe que *PRRT2* interactúa con las proteínas SNAP-25, SYT1 y SYT2 en la membrana presináptica de las neuronas, que participan en la señalización en las células nerviosas. La falta de *PRRT2*, a menudo causada por la vía de degradación del ARN mensajero mediada por mutaciones terminadoras (NMD por sus siglas en inglés), debido a codones de terminación prematura (PTC por sus siglas en inglés) en el transcripto, es el mecanismo molecular común involucrado en la haploinsuficiencia que causa PKD.

Objetivos

El objetivo principal de este trabajo es comprender la etiopatogenia de tres enfermedades neurológicas que afectan a tres familias españolas distintas y estudiar los mecanismos genéticos y moleculares que afectan a estas enfermedades.

Materiales y métodos

Se recogieron muestras de ADN de miembros de tres familias españolas: UGM037 (11 individuos) afectados por EA, UGM471 (7 individuos) afectados por ELP y UGM478 (7 individuos) afectados por PKD. También se dispone de una población de control de individuos sanos y otra de individuos afectados por EA disponible. Las muestras se recogieron con la aprobación de los Comités de Ética en la Investigación de los hospitales correspondientes con el consentimiento informado firmado por los pacientes. La cepa bacteriana utilizada para todas las manipulaciones necesarias fue *Escherichia coli DH5a* y la línea celular humana fue SH-SY5Y.

La secuenciación del exoma completo (WES por sus siglas en inglés) se utilizó como herramienta para identificar variantes dentro del exoma de los pacientes. A través de “base-calling” y análisis de imágenes, alineación de lectura y “SNP calling”, los datos podrían transformarse en una base de datos viable de variantes. Se utilizó el filtrado y la priorización seguida de la validación de secuenciación de Sanger de los resultados para identificar las variantes más interesantes, posiblemente involucradas en la causa de las enfermedades. Para la evaluación de variantes se utilizaron bases de datos públicas como Collaborative Spanish Variant Server (CSVS) o Genome Aggregation Database (gnomAD). La frecuencia de la variante también se verificó mediante PCR específica de alelo (ASPCR por sus siglas en inglés) realizada en las poblaciones de control.

Después de identificar las variantes más interesantes, se realizaron manipulaciones de plásmidos para obtener ADNc, correspondientes a secuencias de los genes con estas variantes, en vectores de expresión específicos. Se utilizó mutagénesis dirigida para la introducción de variantes de un solo nucleótido (SNV por sus siglas en inglés), el epítipo FLAG y sitios de restricción enzimática. La subclonación se utilizó para transferir ADNc específicos a los correspondientes vectores de expresión. El análisis *in silico* se realizó utilizando los servicios web HOPE, I-Mutant 2.0 y ConSurf. De esa forma se pudo establecer la influencia de una variante específica en la función y estabilidad de la proteína. También el grado de conservación de los residuos.

En el caso de genes implicados en la EA, conocer los genotipos de *APOE* y *CRI* es importante y se realizó mediante PCR. En el caso de *APOE*, la PCR fue seguida por restricción enzimática y análisis del patrón de bandas en un gel de agarosa. En el caso de *CRI*, la PCR fue seguida por secuenciación de Sanger o análisis de presencia o ausencia de bandas en un gel de agarosa.

Las células SH-SY5Y se mantuvieron en un medio de crecimiento completo y se utilizó el reactivo de transfección Lipofectamine 2000 para transfectarlas con plásmidos correspondientes a experimentos específicos. Un experimento consistió en comparar los niveles de ARNm y proteínas de genes específicos afectados por las variantes identificadas. En otro, se realizó la misma comparación después de la inhibición de la vía NMD, usando NMDI14. Para medir los niveles de ARNm, se extrajo el ARNm de células SH-SY5Y transfectadas después de aproximadamente 48h de incubación y se obtuvo el ADNc mediante retrotranscripción. Estos se utilizaron en PCR cuantitativa (qPCR por sus siglas en inglés) con cebadores diseñados específicamente para amplificar el producto de la transcripción reversa con controles apropiados. Para medir los niveles de proteína, las proteínas se extrajeron de las células SH-SY5Y transfectadas después de aproximadamente 48h de incubación y se usaron para el análisis por Western Blot (WB). El gen constitutivo de actina se utilizó como control.

Para el análisis funcional de las distintas variantes de ADPRH, las proteínas se extrajeron de las células SH-SY5Y transfectadas y se sometieron a purificación utilizando una columna de afinidad de proteínas marcadas con GST. Después de la diálisis y concentración de la muestra, las proteínas podrían usarse para un ensayo de actividad. Esto implicó la ADP-ribosilación del sustrato por la toxina del cólera (CT) seguida de la hidrólisis de este enlace por ADPRH y sus variantes. Las muestras se cargaron luego en una columna de cromatografía líquida de ultra rendimiento (UPLC por sus siglas en inglés) que permitió la separación, identificación y cuantificación de los componentes basándose en los datos de un detector que mide la absorbancia a 260 nm. También se utilizó co-inmunoprecipitación con extractos de proteínas de células SH-SY5Y cotransfectadas con ADPRH y ADPRHL1 para determinar si tiene lugar la interacción proteína-proteína.

El análisis estadístico se realizó utilizando la prueba t de Student y ANOVA unidireccional o bidireccional para comparar las diferencias entre las medias de los grupos, con una prueba *post hoc* de Tukey de comparaciones múltiples de medias. Se utilizó el software R para realizar las pruebas.

Resultados y discusión

La filtración, priorización y validación mediante secuenciación Sanger identificó la variante rs764542666 en el gen *CR1* que codifica un PTC c.C406T p.R136* ($CR1^{R136*}$) como la causa probable de EA en la familia UGM037. Los datos clínicos sugirieron LOAD

en la familia, que estaría más en consonancia con los factores de riesgo que con las mutaciones causales. Los miembros de la familia poseían el alelo *APOE* $\epsilon 4$, sin embargo, sorprendentemente, el análisis de la familia reveló que un miembro con un genotipo *APOE* $\epsilon 4/\epsilon 4$ y una edad de 88 años no presentaba alteración cognitiva alguna. La variante $CR1^{R136*}$ segrega con la enfermedad en el pedigrí, y no aparece en ninguno de los miembros sin EA. Estudios anteriores mostraron que la isoforma $CR1^{*2}$ de *CR1*, asociada con un mayor riesgo de LOAD, se expresaba menos que $CR1^{*1}$. Según esto, el transcrito de $CR1^{R136*}$ podría causar un efecto similar al activar el NMD y causar sí haploinsuficiencia. El estudio de los niveles de ARNm y proteínas reveló que $CR1^{R136*}$ se expresa en niveles más bajos que el tipo silvestre. Estos niveles aumentaron significativamente después del tratamiento con inhibidor de NMD, lo que sugiere la participación de esta vía en la degradación del transcrito $CR1^{R136*}$. El genotipado de la isoforma $CR1^{*2}$ y rs3818361 en *CR1* en muestras de la familia UGM037, mostró que ninguno de ellos tenía isoforma $CR1^{*2}$ y rs3818361 no se segregó con la enfermedad, no estando presente en ninguna de las muestras de pacientes. Ambos fueron genotipados, ya que se describieron previamente como asociados con la EA. La isoforma $CR1^{*2}$, como se mencionó anteriormente, se expresa a niveles más bajos que $CR1^{*1}$, por lo que posiblemente afecte un aclaramiento de $A\beta$ más bajo y una desregulación del sistema del complemento. El rs3818361 fue genotipado para verificar si puede ser el verdadero culpable de la enfermedad, estando en desequilibrio de ligamiento con $CR1^{R136*}$. Los hijos de los pacientes y el miembro de la familia sano, no fueron diagnosticados con EA, sin embargo, sus muestras de ADN se recolectaron entre las edades de 50 y 57. Pueden estar en riesgo de desarrollar la enfermedad a una edad posterior, especialmente porque todos excepto uno de ellos tenía una copia del alelo *APOE* $\epsilon 4$. Además, tres de siete de ellos tenían la variante $CR1^{R136*}$. Aunque la sobreexpresión *in vitro* de una proteína en un modelo celular tiene sus limitaciones, el mecanismo molecular detrás de la EA en la familia UGM037 parece ser la haploinsuficiencia causada por la destrucción del transcrito $CR1^{R136*}$ por la vía NMD. Son necesarios más estudios funcionales sobre los efectos de $CR1^{R136*}$ y un seguimiento futuro de los miembros más jóvenes de la familia. Según mi conocimiento, $CR1^{R136*}$ (rs764542666) sería la primera mutación causante de EA conocida en el gen *CR1*.

Los datos clínicos disponibles respaldaron el diagnóstico de ELP en miembros de la familia UGM471. A través de la priorización y filtración de datos de WES y la validación por secuenciación de Sanger, se identificaron dos variantes que posiblemente podrían

causar los síntomas en la familia. Uno era una nueva variante c.G884C p.R295P en *ADPRH* (*ADPRH*^{R295P}) y el otro era una mutación c.T497C p.L166P previamente descrita en *PSENI* (rs63750265; *PSENI*^{L166P}). Este último fue descubierto recientemente dentro del genoma de la familia UGM471, por lo que se discutirán cronológicamente, a partir del descubrimiento.

La *ADPRH* es una proteína ubicua que se encuentra en el citoplasma de ratones y humanos. *ADPRH* hidroliza el enlace *N*-glicosídico entre la arginina y la ADP-ribosa, escindiendo la ADP-ribosa del sustrato en el ciclo de ADP-ribosilación. La reacción resulta ser específica para sustratos mono-ADP-ribosilados, debido a la estructura de la proteína. El ciclo de ribosilación de ADP es muy importante para la regulación celular. El variante *ADPRH*^{R295P} se prevé que sea perjudicial por SIFT y Polyphen. No se encontró en ninguna base de datos pública y la secuenciación de Sanger confirmó su segregación con la enfermedad en el pedigrí de la familia UGM471. Sin embargo, dos individuos sanos de esta familia también portaban *ADPRH*^{R295P}. La evaluación de la frecuencia de *ADPRH*^{R295P} por ASPCR en la población de control, mostró que ninguno de los 196 individuos que no eran afectados por ELP eran portadores. Se encontró que la arginina en la posición 295 en *ADPRH* era un residuo conservado y un cambio a una prolina podría afectar la función o estabilidad de la proteína de manera significativa. La reevaluación de los datos de WES mostró un parálogo de *ADPRH*, *ADPRHL1* afectado por una variante *ADPRHL1*^{L294R}, que también se predice que es perjudicial. Esta variante se segrega con la enfermedad en la familia, no estando presente en los individuos sanos. Se especula que *ADPRH* junto con *ADPRHL1* están implicados en el ensamblaje de filamentos de actina y la modulación de la polimerización de actina puede estar implicada en la interrupción del transporte nucleocitoplasmático que es importante en la patogénesis de la ELA. Así, la variante *ADPRH*^{R295P} puede tener una penetrancia baja, provocando la enfermedad solo en ciertos miembros de la familia, o puede ir acompañada de la variante *ADPRHL1*^{L294R}, para afectar a los pacientes. La coimmunoprecipitación no mostró una interacción entre las dos proteínas, sin embargo, pueden funcionar al unísono sin una interacción directa. El ensayo de actividad para *ADPRH*^{R295P}, donde se usó CT para ADP-ribosilar un sustrato y *ADPRH* de tipo silvestre (*ADPRH*^{WT}) y sus variantes (*ADPRH*^{R295P} y *ADPRH*^{D55A/D56A}) se usaron para hidrolizar la ADP-ribosa, mostró que *ADPRH*^{R295P} tenía una eficiencia de actividad similar a la del *ADPRH*^{WT}. La variante *ADPRH*^{D55A/D56A}, en la que se cambiaron los residuos esenciales del sitio activo, no mostró actividad hidrolasa. Aunque este ensayo no mostró disminución de la actividad de la

variante ADPRH^{R295P}, no tuvo en cuenta el posible efecto desestabilizador de la variante. El ensayo RT-qPCR para medir los niveles de ARNm de esta variante mostró que no variaban del tipo silvestre, sin embargo, los niveles de proteína medidos con WB demostraron ser significativamente más bajos que los del ADPRH^{WT} o para otras variantes examinadas (ADPRH^{R295Q}, ADPRH^{D55A/D56A}). Solo la variante ADPRH^{R295*} que se ha descrito previamente, demostró tener niveles de ARNm muy bajos y no pudo detectarse ninguna proteína. Esto sugiere que puede verse afectado por la vía NMD. Si bien es difícil decir si la variante de ADPRH^{R295P} afecta la enfermedad en la familia UGM471, es importante conocer su fuerte efecto sobre la estabilidad de la proteína para el estudio adicional de la mono-ADP-ribosilación poco estudiada.

La mutación c.T497 C p.L166P en *PSEN1* (rs63750265) se descubrió recientemente en la familia UGM471. Aunque las mutaciones en *PSEN1* comúnmente causan EOAD, ha habido informes que asocian *PSEN1* con ELP y ELA. PSEN1^{A431E} y PSEN1^{L381V} son dos ejemplos de mutaciones asociadas con ELP y EA, que provocan un inicio temprano alrededor de los 40 años de edad. Se sabe que la mutación PSEN1^{L166P} causa EA y paraparesia espástica a una edad temprana, siendo el primer caso encontrado en una niña de 15 años. También se encontró que otras mutaciones en la posición L166 en *PSEN1* estaban asociadas con la EA (PSEN1^{L166V}, PSEN1^{L166R} (rs63750265), PSEN1^{L166H} (rs63750265) y PSEN1^{L166del} (rs63751458)). La mayoría de ellos se asociaron con deterioro cognitivo, aunque algunos se asociaron con síntomas motores. La mutación PSEN1^{L166P} causa pérdida parcial de función de escisión de la γ -secretasa y aumenta la proporción A β ₄₂/A β ₄₀ mediante la reducción de los niveles de A β ₄₀. Además, PSEN1 funciona formando canales de fuga de Ca²⁺ del retículo endoplásmico (RE), y la mutación PSEN1^{L166P} interrumpe esa función. El descubrimiento reciente de la mutación PSEN1^{L166P} en la familia UGM471 no permitió una investigación más exhaustiva de sus efectos, pero se confirmó que se segrega con la enfermedad en el pedigrí. La naturaleza agresiva de esta mutación, la edad temprana de aparición y los síntomas motores, sugieren fuertemente que es PSEN1^{L166P} el que causa la enfermedad en la familia. Si bien no está asociado con ELA ni ELP, los pacientes también pueden haber sido afectados por otros factores ambientales o genéticos para producir este fenotipo. Al mismo tiempo, existen informes de pacientes con ELP con mutaciones en *PSEN1*. O el efecto de las mutaciones de *PSEN1* es lo suficientemente heterogéneo como para causar estos diferentes trastornos, los trastornos están mucho más relacionados debido a los

mecanismos moleculares que los causan, o el efecto de la mutación de *PSENI* se ve afectado por la interacción con otras proteínas. Ciertamente, los genes comunes relacionados con la neurodegeneración deben considerarse al identificar una posible causa de una enfermedad, independientemente de la enfermedad con la que estén asociados. Los pacientes con ELP, ELA o paraparesia espástica deben ser investigados para detectar mutaciones en *PSENI*.

En PKD, la filtración, priorización y validación de secuenciación de Sanger identificaron una nueva variante c.C316T p.Q106* en *PRRT2* (*PRRT2*^{Q106*}) como la causa probable de la enfermedad en la familia UGM478. Los datos clínicos de la familia apoyan el diagnóstico de PKD. La evaluación de los resultados de WES conduce a una fuerte sospecha de *PRRT2*^{Q106*} como mutación causante de la enfermedad. Las mutaciones en *PRRT2* son la única causa conocida de PKD hasta el momento y, a menudo, son mutaciones truncadas que conducen a una haploinsuficiencia debido a la destrucción del transcripto portador de PTC por la vía NMD. *PRRT2*^{Q106*} no se encontró en ninguna base de datos, por lo que los únicos datos de frecuencia se obtuvieron a través de nuestro ensayo ASPCR para una población de control de 192 muestras. Ninguno de ellos portaba esta variante, sin embargo, se segregó con la enfermedad en el pedigrí, no estando presente en ninguno de los miembros de la familia que no tenían PKD. De acuerdo con los resultados previos, se encontraron niveles de ARNm significativamente más bajos para las variantes con PTC (*PRRT2*^{Q106*}, *PRRT2*^{Q163*}, *PRRT2*^{Q250*}), en comparación con el tipo silvestre (*PRRT2*^{WT}), en un modelo celular con *PRRT2* y sus variantes sobreexpresadas. *PRRT2*^{Q163*} y *PRRT2*^{Q250*} se eligieron como controles, siendo variantes descritas previamente en *PRRT2*. La inhibición de la ruta NMD, por el tratamiento de células SH-SY5Y transfectadas con NMDI14, aumentó los niveles de ARNm para *PRRT2*^{Q106*} y *PRRT2*^{Q163*} y mostró una tendencia creciente para *PRRT2*^{Q250*}. Esto sugiere que la vía de la NMD puede ser la culpable de la desintegración del ARNm provocada por las PTC en las variantes estudiadas. Curiosamente, los niveles de proteína de la nueva variante *PRRT2*^{Q106*} eran indetectables con WB antes y después del tratamiento con NMDI14, mientras que las otras variantes estudiadas (*PRRT2*^{Q163*} y *PRRT2*^{Q250*}) tenían niveles de proteína significativamente más bajos que el *PRRT2*^{WT} antes del tratamiento y niveles aumentados después. Esto puede deberse a que las regiones ricas en prolina son importantes para la estabilidad de la proteína, ya que *PRRT2*^{Q106*} tiene una PTC antes de esa región y *PRRT2*^{Q163*} y *PRRT2*^{Q250*} en el medio y después de ella. Estos resultados sugieren que la nueva variante *PRRT2*^{Q106*} es probablemente la causa de PKD

en la familia española UGM478. Los mecanismos moleculares responsables de la aflicción pueden ser la vía NMD que causa la degradación del transcrito que conduce a la haploinsuficiencia. La falta de *PRRT2* a su vez provoca hiperexcitabilidad a través de la liberación desregulada de neurotransmisores y la hiperactividad de los canales de Na^+ .

Mecanismos moleculares patológicos comunes, o relacionados, pueden afectar diversos trastornos neurológicos tradicionalmente considerados como no relacionados, en la intrincada red del sistema nervioso. En este trabajo he esbozado algunos de esos supuestos mecanismos. Variantes de un solo nucleótido que pueden afectar a diferentes fenotipos a través de la pérdida parcial de función debido a la desestabilización de proteínas o la haploinsuficiencia debida a NMD.

El número estimado de genes haploinsuficientes humanos es 12,443 de aproximadamente 22,000. Si bien el número total de genes humanos es un tema de debate y estudio adicional, su estimación indica que podemos esperar un gran número de genes en los que el efecto de la dosis del nivel de proteína puede ser esencial. La haploinsuficiencia es importante en los trastornos neurológicos. Recientemente se descubrió que *C9orf72* es haploinsuficiente en ELA/DFT debido a la expansión repetida de GGGGCC.

En este trabajo, postulo que *CRI* y *PRRT2* son haploinsuficientes en familias españolas con EA y PKD respectivamente. Si bien está establecido que la mayoría de las mutaciones en *PRRT2* conducen a la pérdida de función y a la haploinsuficiencia, según mi conocimiento, no existen tales informes sobre *CRI*. La haploinsuficiencia, por tanto, surge como un factor común entre estas y otras enfermedades neurológicas. Además, el mecanismo molecular detrás de la haploinsuficiencia relacionada con *CRI* y *PRRT2* parece ser la NMD provocada por SNV que codifican PTC, lo que demuestra un mecanismo molecular común en distintas enfermedades neurológicas.

La pérdida de función está estrictamente relacionada con la haploinsuficiencia, que es un fenotipo dominante en organismos heterocigotos para tales alelos. Aunque no se encontró que la variante *PSEN1*^{L166P} (rs63750265) causara haploinsuficiencia, afecta una pérdida parcial de la función de escisión de la γ -secretasa y la función del canal de fuga del Ca^{2+} del RE. Se muestra que la variante *ADPRH*^{R295P}, identificada en la misma familia, desestabiliza significativamente la proteína, afectando drásticamente sus niveles. Esto, a su vez, puede obstaculizar su función. Queda por ver si la variante *ADPRH*^{R295P} en heterocigosis es, de hecho, deletérea, dependiendo de su tolerancia a la dosis de proteína disminuida. Sin embargo, la pérdida de función, ya sea

total o parcial, engloba los mecanismos moleculares subyacentes de los SNV descritos en este trabajo, que contribuyen a enfermedades neurológicas independientes.

Al discutir los mecanismos moleculares en EA y ELP en las dos familias españolas, y la participación de las variantes CR1^{R136*} (rs764542666) y PSEN1^{L166P} (rs63750265) en la patogénesis de las enfermedades, es importante no omitir otros factores posiblemente contribuyentes. Si bien CR1^{R136*} puede ser una mutación causal en la familia, sus miembros tenían una alta incidencia de *APOE* ϵ 4. En familias con mutaciones de *APP*, la incidencia de *APOE* ϵ 4 se relacionó con una edad de inicio más temprana, mientras que la incidencia de *APOE* ϵ 2, con una edad de inicio más tardía, con respecto a *APOE* ϵ 3. Curiosamente, la variante PSEN1^{E318G} está relacionada con un mayor riesgo de EA, que depende de *APOE* ϵ 4. Si bien, por lo demás, PSEN1^{E318G} se consideró no patógeno, su interacción con *APOE* ϵ 4 aumentó la deposición de A β , provocando un deterioro cognitivo más rápido y una neurodegeneración. Así, si bien portar la variante CR1^{R136*} puede ser suficiente para desarrollar EA, también es probable que los miembros de la familia española estudiados estuvieran afectados únicamente por el factor de riesgo *APOE* ϵ 4 o una combinación de ambos.

De manera similar, otros factores, ya sean ambientales o genéticos, pueden afectar los síntomas desarrollados por la familia con ELP. Aunque PSEN1^{L166P} parece ser responsable del fenotipo experimentado por los pacientes, sus síntomas difieren de las características más canónicas de la EA relacionadas con esta variante. Esta expresión divergente de la enfermedad puede deberse a la pleiotropía del gen *PSEN1*, sin embargo, también puede deberse a otros factores contribuyentes. Se ha descubierto que los factores ambientales juegan un papel importante en la ELA y no se pueden ignorar en una enfermedad familiar. Aquí propongo un factor genético que puede contribuir a los distintos síntomas que experimentan los miembros de esta familia. La nueva variante ADPRH^{R295P} puede tener un efecto sobre el desarrollo de la enfermedad, desestabilizando los filamentos de actina en presencia de la variante ADPRHL1^{L294R}, provocando un fenotipo más cercano a ELP, junto con la mutación agresiva PSEN1^{L166P}. La complejidad de las enfermedades neurológicas proviene, en parte, de la naturaleza acumulativa de los defectos que las provocan, por lo que siempre es fundamental buscar otro factor que pueda contribuir al fenotipo observado. Se necesitan más estudios sobre los efectos que las variantes de ADPRH pueden tener sobre las enfermedades neurológicas, ya que contribuye a la mono-ADP-ribosilación aún poco conocida, pero muy importante.

Las limitaciones generales de la sobreexpresión *in vitro* de una proteína en un modelo celular se aplican a todo el estudio. Diferencias entre el modelo celular y las células correspondientes en el organismo, problemas para establecer un microambiente apropiado, como interacciones con otras células, o el hecho de que la proteína se sobreexpresa artificialmente en cantidades naturalmente no disponibles. Pueden ser necesarios más estudios funcionales para todas las variantes descritas.

Para concluir, la variante rs764542666 en el gen *CR1* que codifica un PTC c.C406T p.R136* es la causa probable de la EA en una familia española UGM037, basado en WES y estudio de la expresión genética. La haploinsuficiencia provocada por la vía NMD de *CR1* es el probable mecanismo molecular detrás de la enfermedad. La variante rs764542666 es probablemente la primera mutación causante de EA conocida en *CR1*, lo que fomenta la investigación de las raras variantes truncadas en este gen.

El mutante rs63750265 en el gen *PSEN1* que codifica una mutación c.T497C p.L166P es la causa probable de ELP en una familia española UGM471, según el estudio WES, análisis de segregación y conocimientos previos, lo que plantea dudas sobre los efectos pleiotrópicos de la mutación. Los mecanismos moleculares detrás del mutante rs63750265 que causa la enfermedad en la familia UGM471 son probablemente la pérdida de la actividad de la γ -secretasa, el aumento de la relación $A\beta_{42}/A\beta_{40}$ y el deterioro de la función del canal de fuga de Ca^{+2} del RE. La nueva variante en el gen *ADPRH* que codifica una variante c.G884C p.R295P desestabiliza fuertemente la proteína sin afectar su función, arrojando luz sobre el estudio de la mono-ADP-ribosilación.

La nueva variante en el gen *PRRT2* que codifica un PTC c.C316T p.Q106* es la causa probable de PKD en una familia española UGM037, según el estudio de expresión genética y WES. Este trabajo apoya la hipótesis de que la vía de NMD provoca la haploinsuficiencia de *PRRT2*, como mecanismo molecular detrás de la PKD.

TABLE OF CONTENTS

TABLE OF CONTENTS.....	1
LIST OF TABLES	8
LIST OF FIGURES	11
ABBREVIATIONS	14
1. INTRODUCTION	21
1.1. Molecular mechanisms of neurological diseases	21
1.2. Alzheimer’s disease.....	21
1.2.1. Background.....	21
1.2.2. Epidemiology.....	22
1.2.2.1. Prevalence	23
1.2.2.2. Incidence	24
1.2.2.3. Alzheimer’s disease epidemiology in Spain	25
1.2.3. Clinical features and treatment	27
1.2.4. Alzheimer’s disease pathology	28
1.2.4.1. Amyloid plaques	28
1.2.4.2. Neurofibrillary tangles	29
1.2.4.3. Progression of Alzheimer’s disease pathology through the brain.....	30
1.2.5. Etiology.....	32
1.2.5.1. Environmental factors	32
1.2.5.2. Genetic basis of Alzheimer’s disease.....	32
1.2.6. Complement system.....	37
1.2.6.1. Complement system in Alzheimer’s disease pathogenesis	39
1.2.6.2. Complement receptor 1	40
1.3. Amyotrophic lateral sclerosis and primary lateral sclerosis.....	43

1.3.1.	Amyotrophic lateral sclerosis background.....	43
1.3.2.	Primary lateral sclerosis background	44
1.3.3.	Epidemiology	45
1.3.3.1.	Prevalence	45
1.3.3.2.	Incidence	46
1.3.3.3.	Amyotrophic lateral sclerosis epidemiology in Spain.....	47
1.3.4.	Clinical features and treatment	48
1.3.5.	Amyotrophic lateral sclerosis pathology	50
1.3.5.1.	RNA processing	51
1.3.5.2.	Mitochondrial dysfunction	51
1.3.5.3.	Glutamate excitotoxicity	52
1.3.5.4.	Oxidative stress	52
1.3.5.5.	Protein aggregation	53
1.3.5.6.	Neurofilaments accumulation	53
1.3.5.7.	Neuroinflammation	54
1.3.6.	Etiology.....	55
1.3.6.1.	Environmental factors	55
1.3.6.2.	Genetic basis of amyotrophic lateral sclerosis and primary lateral sclerosis.....	55
1.3.7.	ADP-ribosylhydrolases and ADP-ribosyltransferases.....	61
1.4.	Paroxysmal kinesigenic dyskinesia.....	64
1.4.1.	Background.....	64
1.4.2.	Epidemiology.....	65
1.4.3.	Paroxysmal dyskinesias	66
1.4.4.	Clinical features and treatment	67
1.4.5.	<i>PRRT2</i> and paroxysmal kinesigenic dyskinesia	68
1.4.5.1.	<i>PRRT2</i> gene.....	71

1.4.6.	Other genetic risk factors for paroxysmal kinesigenic dyskinesia.....	72
1.4.7.	Nonsense-mediated mRNA decay (NMD) pathway.....	73
2.	OBJECTIVES.....	75
3.	MATERIALS AND METHODS	77
3.1.	Nomenclature	77
3.2.	Subjects	77
3.2.1.	Family UGM037	77
3.2.2.	Family UGM471	78
3.2.3.	Family UGM478.....	79
3.2.4.	Control population	80
3.3.	Materials.....	80
3.3.1.	Bacterial strain	80
3.3.2.	Cell line.....	81
3.3.3.	Buffers and media solutions.....	81
3.3.3.1.	Buffers and media used for cultivation of bacteria and genetic engineering	81
3.3.3.2.	Buffers and media used for protein biochemistry	83
3.3.3.3.	Buffers and media used with SH-SY5Y neuroblastoma cells.....	85
3.3.4.	Expression vectors	86
3.3.5.	Antibodies	88
3.4.	Methods.....	88
3.4.1.	Whole-exome sequencing.....	88
3.4.1.1.	Base-calling and image analysis	89
3.4.1.2.	Read alignment.....	89
3.4.1.3.	SNP calling.....	90
3.4.1.4.	Results filtering and identification of interesting variants	90
3.4.1.5.	Public databases	91

3.4.2.	Sanger sequencing confirmation of WES results.....	91
3.4.3.	Allele specific PCR.....	92
3.4.4.	Plasmid manipulations	93
3.4.4.1.	Site-directed mutagenesis.....	93
3.4.4.2.	Subcloning.....	96
3.4.4.3.	Plasmid constructs.....	99
3.4.5.	<i>In silico</i> analysis.....	99
3.4.6.	Genotyping.....	100
3.4.6.1.	Genotyping apolipoprotein E gene polymorphism	100
3.4.6.2.	Genotyping <i>CRI</i> gene polymorphism	102
3.4.7.	SH-SY5Y cell culture maintenance and transfection	104
3.4.7.1.	Maintenance	104
3.4.7.2.	Transfection.....	105
3.4.8.	Nonsense-mediated mRNA decay (NMD) pathway inhibition analysis	106
3.4.9.	mRNA levels analysis.....	107
3.4.9.1.	Cellular RNA extraction and quantification.....	107
3.4.9.2.	DNase I treatment and reverse transcription	107
3.4.9.3.	Quantitative PCR (qPCR)	108
3.4.10.	Protein expression analysis	109
3.4.10.1.	Protein extraction and quantification	109
3.4.10.2.	Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS- PAGE).....	109
3.4.10.3.	Western Blot (WB).....	109
3.4.10.4.	Western blot membrane stripping	110
3.4.11.	Extraction and purification of protein for activity assay	110
3.4.11.1.	Extraction of glutathione S-transferase (GST)-tagged protein	110
3.4.11.2.	Purification of GST-tagged protein.....	111

3.4.11.3.	Sample dialysis.....	112
3.4.11.4.	Protein concentration.....	112
3.4.12.	Activity assay for ADPRH with cholera toxin	112
3.4.13.	Co-immunoprecipitation (co-IP).....	114
3.4.14.	Statistical analysis.....	115
4.	RESULTS.....	117
4.1.	Study of etiopathogenesis of AD in family UGM037.....	117
4.1.1.	Clinical data	117
4.1.2.	WES data	117
4.1.2.1.	Base-calling and image analysis raw data.....	117
4.1.2.2.	Read alignment results	118
4.1.2.3.	SNP calling results	118
4.1.3.	WES results filtration and identification of interesting variants.....	119
4.1.4.	Sanger sequencing verification of selected variants	123
4.1.5.	Frequency of the rs764542666 variant in control population.....	125
4.1.6.	Genotyping <i>APOE</i> and <i>CRI</i> polymorphisms.....	126
4.1.7.	Role of variant rs764542666 (p.R136*) in <i>CRI</i> mRNA expression	127
4.1.8.	Role of variant rs764542666 (p.R136*) in CR1 protein expression	129
4.2.	Study of etiopathogenesis of PLS in family UGM471	131
4.2.1.	Clinical data	131
4.2.2.	WES data	132
4.2.2.1.	Base-calling and image analysis raw data.....	132
4.2.3.	WES results filtration and identification of interesting variants.....	133
4.2.4.	Sanger sequencing verification of selected variants and result assessment.....	135
4.2.5.	Frequency of the ADPRH ^{R295P} variant in control population.....	137
4.2.6.	ADPRH ^{R295P} variant's <i>in silico</i> analysis	138

4.2.6.1.	ConSurf estimation of evolutionary conservation.....	138
4.2.6.2.	HOPE protein structure analysis of ADPRH ^{R295P} variant.....	139
4.2.6.3.	Prediction of protein stability by I-Mutant 2.0 web server	141
4.2.7.	Reassessment of WES results	142
4.2.7.1.	Sanger sequencing verification of variants selected in WES results reassessment	143
4.2.8.	Role of variant ADPRH ^{R295P} in protein activity	144
4.2.9.	Role of variant ADPRH ^{R295P} in mRNA expression.....	149
4.2.10.	Role of variant ADPRH ^{R295P} in protein expression	150
4.2.11.	Protein-protein interaction between ADPRH and ADPRHL1.....	151
4.2.12.	PSEN1 ^{L166P} (rs63750265) mutant and AD	151
4.2.12.1.	Sanger sequencing verification of mutation PSEN1 ^{L166P} (rs63750265).....	152
4.3.	Study of etiopathogenesis of PKD in family UGM478	153
4.3.1.	Clinical data	153
4.3.2.	WES data	154
4.3.2.1.	Base-calling and image analysis raw data.....	154
4.3.2.2.	Read alignment results	154
4.3.2.3.	SNP calling results	155
4.3.3.	WES results filtration and identification of interesting variants.....	156
4.3.4.	Sanger sequencing verification of selected variants	159
4.3.5.	Frequency of the PRRT2 ^{Q106*} variant in control population	161
4.3.6.	Role of variant PRRT2 ^{Q106*} in mRNA expression.....	161
4.3.7.	Role of variant PRRT2 ^{Q106*} in protein expression.....	164
5.	DISCUSSION.....	167
5.1.	Overview	167
5.2.	Etiopathogenesis of AD in family UGM037.....	167

5.2.1. Identification of a variant possibly responsible for AD in a Spanish family.....	168
5.3. Etiopathogenesis of PLS in family UGM471	171
5.3.1. Identification of a variant possibly responsible for PLS in a Spanish family.....	172
5.3.1.1. Variant ADPRH ^{R295P} in Spanish family affected by primary lateral sclerosis.....	172
5.3.1.2. Variant PSEN1 ^{L166P} in Spanish family affected by primary lateral sclerosis.....	175
5.4. Etiopathogenesis of PKD in family UGM478	178
5.4.1. Identification of a variant possibly responsible for PKD in a Spanish family.....	178
5.5. Final remarks.....	180
6. CONCLUSIONS	183
7. REFERENCES	185
Appendix I License agreements.....	221
Appendix II Expression vectors.....	232
Appendix III Commercial kits	236
Appendix IV Oligonucleotide primers.....	237

LIST OF TABLES

- Table 1.1. Frequency of *APOE* genotypes in a US population.
- Table 1.2. Selected genes and their variants associated with Alzheimer's disease.
- Table 1.3. Genes associated with fALS and their inheritance and clinical features.
- Table 1.4. Summary of ADP-ribosylation intricate array of chemical linkages, transferases, hydrolases and their substrates.
- Table 1.5. Classification and information summary of paroxysmal dyskinesias according to Demirkiran and Jankovic (1995).
- Table 1.6. PRRT2 mechanisms of involvement in PKD pathogenesis.
- Table 3.1. Expression vectors used in this study.
- Table 3.2. Primary and secondary antibodies used in this study.
- Table 3.3. Plasmid constructs created through site-directed mutagenesis and subcloning.
- Table 3.4. Interpretation of the *APOE* gene polymorphism genotyping.
- Table 3.5. Flasks and dish used for SH-SY5Y cell culture.
- Table 3.6. Groups of SH-SY5Y cells transfected with corresponding plasmids used in NMD pathway inhibition analysis.
- Table 4.1. Data including diagnosis, age of onset and age of death for family UGM037 affected members and their sibling.
- Table 4.2. WES raw data: reads, base content and phred quality scores for samples 04-046 and 04-048.
- Table 4.3. WES post-alignment statistics for samples 04-046 and 04-048.
- Table 4.4. WES SNP calling summary for samples 04-046 and 04-048.
- Table 4.5. WES results filtered for AD associated genes for samples 04-046 and 04-048.
- Table 4.6. WES results filtered for possibly damaging variants for samples 04-046 and 04-048.
- Table 4.7. Interpretation of the selected variants according to Richards *et al.* (2015) and observed/expected score.
- Table 4.8. Verification of selected variants through Sanger sequencing in family UGM037.
- Table 4.9. General population frequency for rs764542666 with data on specific populations.
- Table 4.10. Number of cases of rs764542666 variant in different groups.

Table 4.11. Data including diagnosis and age of onset for family UGM471 affected members.

Table 4.12. WES raw data for samples 0002, 1001 and 1002.

Table 4.13. WES results filtered for PLS associated genes for samples 0002, 1001 and 1002.

Table 4.14. WES results filtered for possibly damaging variants for samples 0002, 1001 and 1002.

Table 4.15. Interpretation of the selected variants according to the observed/expected score.

Table 4.16. Verification of selected variants through Sanger sequencing in family UGM471.

Table 4.17. Summary of known variants in *ADPRH* at amino acid position 295.

Table 4.18. Number of cases of *ADPRH*^{R295P} variant in different groups.

Table 4.19. I-Mutant 2.0 protein stability predictions for SNVs at position R295 in *ADPRH*.

Table 4.20. WES results filtered for possibly damaging variants in genes associated with *ADPRH* pathways for samples 0002, 1001 and 1002.

Table 4.21. Verification of selected variants through Sanger sequencing in family UGM471.

Table 4.22. Control runs of the *ADPRH* activity assay.

Table 4.23. Experimental runs of the *ADPRH* activity assay.

Table 4.24. WES results for *PSEN1*^{L166P} (rs63750265) mutation for samples 0002, 1001 and 1002.

Table 4.25. Verification of mutation *PSEN1*^{L166P} (rs63750265) through Sanger sequencing in family UGM471.

Table 4.26. Family UGM478 patient data.

Table 4.27. WES raw data: reads, base content and phred quality scores for samples 1005, 2005 and 2007.

Table 4.28. WES post-alignment statistics for samples 1005, 2005 and 2007.

Table 4.29. WES SNP calling summary for samples 1005, 2005 and 2007.

Table 4.30. WES results filtered for PKD and PxD associated genes for samples 1005, 2005 and 2007.

Table 4.31. WES results filtered for possibly damaging variants for samples 1005, 2005 and 2007.

Table 4.32. Interpretation of the selected variants according to Richards *et al.* (2015) and observed/expected score.

Table 4.33. Verification of selected variants through Sanger sequencing in family UGM478.

Table 4.34. Number of cases of PRRT2^{Q106*} variant in different groups.

LIST OF FIGURES

Figure 1.1. Prevalence rates of Alzheimer's disease for different age groups.

Figure 1.2. Annual incidence rate (per 100 person-years) of AD (modified from Mayeux and Stern, 2012).

Figure 1.3. Incidence rate (per 1000 person-years) of Alzheimer's disease in a Spanish cohort from Zaragoza (data from Lobo et al., 2011).

Figure 1.4. Comparison of incidence rates (per 1000 person-years) of Alzheimer's disease between sexes of a Spanish cohort from Zaragoza (data from Lobo et al., 2011).

Figure 1.5. Hallmark lesions of Alzheimer's disease visualised by modified Bielschowsky Silver Stain (from Selkoe *et al.*, 1991).

Figure 1.6. Comparison of normal state and Alzheimer's disease state, with amyloid β aggregates and neurofibrillary tangles formation as the main hallmarks of the disease (from Gireud *et al.*, 2014).

Figure 1.7. Spread of neurofibrillary pathology in Alzheimer's disease (modified from Smith, 2002).

Figure 1.8. Pathways of the complement cascade.

Figure 1.9. Schematic representation of CR1 isoform structures.

Figure 1.10. Incidence of amyotrophic lateral sclerosis per 100,000 person-years.

Figure 1.11. Amyotrophic lateral sclerosis pathogenesis mechanisms.

Figure 1.12. Approximate contribution of mutations in sALS and fALS. The sALS representing approximately 90 % and the fALS 10 % of all ALS patients (data from Morgan and Orell, 2016).

Figure 1.13. (A) Structure of *PRRT2* gene and protein. (B) Membrane topology of PRRT2 protein (modified with permission from Valtorta *et al.*, 2016).

Figure 1.14. Map of nonsense/frameshift and missense mutations in proline-rich transmembrane protein 2 (PRRT2) with regards to their protein domain location (modified with permission from Valtorta *et al.*, 2016).

Figure 3.1. Family tree of UGM037 family with the three affected members denoted by coloured in shapes.

Figure 3.2. Family tree of UGM471 family with the three affected members denoted by coloured in shapes.

Figure 3.3. Family tree of UGM478 family with the five affected members denoted by coloured in shapes.

Figure 3.4. SH-SY5Y (ATCC® CRL-2266™) neuroblastoma cells.

Figure 4.1. UGM037 family tree with the three AD affected members denoted by coloured in shapes. Genotypes for rs764542666 variant are included.

Figure 4.2. UGM037 family tree with the three AD affected members denoted by coloured in shapes. Genotypes for rs764542666 and rs3818361 variants as well as APOE and CR1*2 isoform are included.

Figure 4.3. RT-qPCR results for mRNA levels of CR1^{WT} and CR1^{R136*} variant, overexpressed in SH-SY5Y cells.

Figure 4.4. RT-qPCR results for mRNA levels of CR1^{WT} and CR1^{R136*} variant overexpressed in SH-SY5Y cells with and without NMD pathway inhibition.

Figure 4.5. Ratio of mRNA levels change after treatment with NMDI14 for CR1^{WT} and CR1^{R136*} variant overexpressed in SH-SY5Y cells.

Figure 4.6. WB results for protein levels of CR1^{WT} and CR1^{R136*} variant, overexpressed in SH-SY5Y cells.

Figure 4.7. WB results for protein levels of CR1^{WT} and CR1^{R136*} variant overexpressed in SH-SY5Y cells with and without NMD pathway inhibition.

Figure 4.8. Ratio of protein levels increase after treatment with NMDI14 for CR1^{WT} and CR1^{R136*} variant overexpressed in SH-SY5Y cells.

Figure 4.9. UGM471 family tree with the three PLS affected members denoted by coloured in shapes. Genotypes for ADPRH^{R295P} variant are included.

Figure 4.10. Multiple sequence alignment of a segment of the ADPRH protein from seven different species created using MultAlin (Corpet, 1988; <http://multalin.toulouse.inra.fr/multalin/>).

Figure 4.11. Crystal structure of ADPRH protein created using Mol* (Sehnal *et al.*, 2018) as deposited in the RCSB Protein Data Bank (Berman *et al.*, 2000; rcsb.org) by Liu and Yu (PDB ID: 6IUX; 2019).

Figure 4.12. Close-up of the residue 295 in ADPRH protein.

Figure 4.13. UPLC chromatograms representing three separations to identify ADP-ribosyl-agmatine.

Figure 4.14. Selected UPLC chromatograms representing five separations to measure the relative activity of ADPRH and its variants ADPRH^{R295P} and ADPRH^{D55A/D56A}.

Figure 4.15. ADPRH activity assay with agmatine and CT using UPLC.

Figure 4.16. RT-qPCR results for mRNA levels of ADPRH^{WT}, ADPRH^{R295P}, ADPRH^{R295Q}, ADPRH^{R295*} and ADPRH^{D55A/D56A} overexpressed in SH-SY5Y cells.

Figure 4.17. Western blot results for ADPRH^{WT} and ADPRH^{R295P}, ADPRH^{R295Q}, ADPRH^{R295*} and ADPRH^{D55A/D56A} variants overexpression in SH-SY5Y cells.

Figure 4.18. UGM471 family tree with the three PLS affected members denoted by coloured in shapes. Genotypes for PSEN1^{R295P} (rs63750265) mutant and ADPRH^{R295P} variant are included.

Figure 4.19. UGM478 family tree with the five PKD affected members denoted by coloured in shapes. Genotypes for PRRT2^{Q106*} variant are included.

Figure 4.20. Structure of PRRT2 protein with map of nonsense variants PRRT2^{Q106*}, PRRT2^{Q163*} and PRRT2^{Q250*} with regards to their protein domain location (modified with permission from Valtorta *et al.*, 2016).

Figure 4.21. RT-qPCR results for mRNA levels of PRRT2^{WT} and PRRT2^{Q106*}, PRRT2^{Q163*} and PRRT2^{Q250*} variants overexpressed in SH-SY5Y cells.

Figure 4.22. RT-qPCR results for mRNA levels of PRRT2^{WT} and PRRT2^{Q106*}, PRRT2^{Q163*} and PRRT2^{Q250*} variants overexpressed in SH-SY5Y cells with and without NMD pathway inhibition.

Figure 4.23. Ratio of mRNA levels change after treatment with NMDI14 for PRRT2^{WT} and PRRT2^{Q106*}, PRRT2^{Q163*} and PRRT2^{Q250*} variants overexpressed in SH-SY5Y cells.

Figure 4.24. Western blot results for PRRT2^{WT} and PRRT2^{Q106*}, PRRT2^{Q163*} and PRRT2^{Q250*} variants overexpression in SH-SY5Y cells.

Figure 4.25. Western blot results for PRRT2^{WT} and PRRT2^{Q106*}, PRRT2^{Q163*} and PRRT2^{Q250*} variants overexpression in SH-SY5Y cells.

Figure 4.26. Ratio of protein levels increase after treatment with NMDI14 for PRRT2^{WT} and PRRT2^{Q163*} and PRRT2^{Q250*} variants overexpressed in SH-SY5Y cells.

ABBREVIATIONS

1000G	1000 Genomes Project
AD	Alzheimer's disease
ADNFLE	Autosomal dominant nocturnal frontal lobe epilepsy
ADP	Adenosine diphosphate
AICD	APP-intracellular domain
ALS	Amyotrophic lateral sclerosis
<i>ALS2</i>	Alsin
Alt	Alternate allele
AMPA	α -amino-3-hydroxyl-5-methyl-4-isoxazole-propionate
<i>APOE</i>	Apolipoprotein E
<i>APP</i>	Amyloid precursor protein
ARHs	ADP-ribosylhydrolases
ARTs	ADP-ribosyltransferases
ASPCR	Allele-specific PCR
A β	Amyloid- β
BAM	Binary Alignment/Map
BBB	Blood-brain barrier
BFIE	Benign familial infantile epilepsy
BMAA	β -N-methylamino-l-alanine
bp	Base pairs
BWA	Burrows-Wheeler Alignment
<i>C9orf72</i>	Chromosome 9 open reading frame 72
CBZ	Carbamazepine
ChEI	cholinesterase inhibitors
Chr	Chromosome
CI	Confidence interval
<i>CLCN1</i>	Chloride voltage-gated channel 1
CNS	Central nervous system
Co-IP	Co-immunoprecipitation
<i>CRI</i>	Complement receptor 1
CSF	Cerebrospinal fluid

CSVS	Collaborative Spanish Variant Server
CT	Cholera toxin B subunit
DDG	Delta delta G
ddH ₂ O	double-distilled water
DECID	Decay-inducing complex
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DPR	Dipeptide repeat
DTT	Dithiothreitol
EAATs	Excitatory amino acid transporters
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
EEG	Electroencephalography
EJCs	Exon-exon junction complexes
EMA	European Medicines Agency
EOAD	Early onset Alzheimer's disease
EO-fAD	Early onset familial Alzheimer's disease
ER	Endoplasmic reticulum
eRF	Eukaryotic release factor
ESI	Electrospray ionization
ExAC	Exome Aggregation Consortium
fAD	Familial Alzheimer's disease
fALS	Familial amyotrophic lateral sclerosis
FBS	Fetal bovine serum
FDA	Food and Drug Administration
fPKD	Familial paroxysmal kinesigenic dyskinesia
fPLS	Familial primary lateral sclerosis
FTD	Frontotemporal dementia
<i>FUS</i>	Fused in sarcoma
GATK	Genome Analysis Toolkit
gnomAD	Genome Aggregation Database
GO-ESP	NHLBI GO Exome Sequencing Project
GPI	Glycosylphosphatidylinositol

GST	Glutathione S-transferase
GWAS	Genome-wide association studies
HCS	HiSeq Control Software
HD	Huntington disease
Het	Heterozygous
HGNC	HUGO Gene Nomenclature Committee
Hom	Homozygous
HQ	High quality
HSP	Hereditary spastic paraparesis
IBMCP	Institute for Plant Molecular and Cell Biology
ICCA	Infantile convulsions and paroxysmal choreoathetosis
INCS	Infantile non-convulsive seizures
IPTG	Isopropyl β -D-1-thiogalactopyranoside
LB	Lysogeny broth
LHRs	Long homologous repeats
LMNs	Lower motor neurons
LOAD	Late onset Alzheimer's disease
LOD	Logarithm of odds
LO-fAD	Late onset familial Alzheimer's disease
MAC	Membrane attack complex
<i>MAPT</i>	microtubule associated protein tau
MARylation	Mono-(ADP-ribosyl)ation
MBL	Mannose-binding lectin
MC	Myotonia congenita
MND	Motor neuron disease
MR-1	Myofibrillogenesis regulator 1
MRI	Magnetic resonance imaging
mRNA	Messenger RNA
NC	Nocturnal convulsions
NEAAs	Non-essential amino acids
<i>NEFH</i>	Heavy chain neurofilament
NF-H	Neurofilament heavy subunit
NF-L	Neurofilament light subunit

NF-M	Neurofilament medium subunit
NFs	Neurofilaments
NFTs	Neurofibrillary tangles
NGS	Next-generation sequencing
NMD	Nonsense-mediated mRNA decay
NMDA	<i>N</i> -methyl-D-aspartate
OAADPr	O-Acetyl-ADP-ribose
OD	Optical density
<i>OPTN</i>	Optineurin
ORFs	Open reading frames
<i>PARP1</i>	PAR polymerase 1
<i>PARP2</i>	PAR polymerase 2
PARylation	Poly-(ADP-ribosyl)ation
<i>PAXIP1</i>	PAX Interacting Protein 1
PBS	Phosphate buffered saline
PCB	PreScission cleavage buffer
PCR	Polymerase chain reaction
PD	Parkinson's disease
PDB	Protein Data Bank
PDC	Pyruvate dehydrogenase complex
PED	Paroxysmal exercise-induced dyskinesia
PEG	Percutaneous endoscopic gastrostomy
PET	Positron-emission tomography
PfEMP1	<i>Plasmodium falciparum</i> erythrocyte membrane protein 1
PHD	Paroxysmal hypnogenic dyskinesia
PIPLC	Phosphatidylinositol-specific phospholipase C
PKD	Paroxysmal kinesigenic dyskinesia
PLS	Primary lateral sclerosis
PMA	Progressive muscular atrophy
PNKD	Paroxysmal non-kinesigenic dyskinesia
PNS	Peripheral nervous system
Polyphen	Polymorphism Phenotyping
Pos	Position

<i>PRRT2</i>	Proline-rich transmembrane protein 2
<i>PSEN1</i>	Presenilin 1
<i>PSEN2</i>	Presenilin 2
<i>PSENEN</i>	Presenilin Enhancer Protein 2
PTCs	Premature termination codons
PTM	Posttranslational modification
PxDs	Paroxysmal dyskinesias
RAN	Repeat-associated non-AUG
RBPs	RNA binding proteins
Ref	Reference allele
RI	Reliability index
RIPA	Radioimmunoprecipitation assay
RNA	Ribonucleic acid
ROS	Reactive oxygen species
RT	Retention time
RTA	Real Time Analysis
RT-qPCR	reverse transcription quantitative PCR
sALS	Sporadic amyotrophic lateral sclerosis
SAM	Sequence Alignment/Map
sAPP α	Soluble amyloid precursor protein α
sAPP β	Soluble amyloid precursor protein β
SBS	Sequencing by synthesis
sCR1	Soluble complement receptor 1
SCRs	Short consensus repeats
SD	Standard deviation
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate–polyacrylamide gel electrophoresis
SEM	Standard error of the mean
SIFT	Sorting Intolerant From Tolerant
<i>SLC2A1</i>	Solute carrier family 2 member 1
<i>SNAP-25</i>	Synaptosomal nerve-associated protein 25
SNP	Single nucleotide polymorphism
SNV	Single nucleotide variant

SOB	Super optimal broth
SOC	SOB with catabolite repression
<i>SOD1</i>	Superoxide dismutase 1
<i>SPG11</i>	Spatacsin
STR	Short tandem repeats
SURF	Surveillance complex
<i>SYT1</i>	Synaptotagmin-1
<i>SYT2</i>	Synaptotagmin-2
<i>TARDBP</i>	TAR DNA binding protein 43
TOPMed	Trans-Omics for Precision Medicine
Ts	Transition
Tv	Transversion
<i>UBQLN2</i>	X chromosome linked ubiquilin-2
UCSC	University of California Santa Cruz
UMNs	Upper motor neurons
<i>UPF1</i>	Up-frameshift 1
UPLC	Ultraperformance liquid chromatography
UPLC-qTOF/MS	UPLC-quadrupole time-of-flight mass spectrometry
UTR	Untranslated region
<i>VAPB</i>	Vehicle-associated membrane protein B
<i>VCP</i>	Valosin-containing protein
WES	Whole-exome sequencing
WGS	Whole-genome sequencing
WT	Wild-type
β -CTFs	β -C-terminal fragments
β -NAD ⁺	β -nicotinamide adenine dinucleotide

1. INTRODUCTION

1.1. Molecular mechanisms of neurological diseases

Neurological disorders, traditionally considered as unrelated, may share common or related pathological molecular mechanisms. The vast, intricate network of the nervous system processes, sends and receives a remarkable amount of information which depends on the workings of many of those mechanisms. Both specific to the nervous system and those common to every cell of the body. The complexity of neurological diseases comes, in part, from a cumulative nature of defects that cause them. This often manifests itself by late onset of the diseases, affecting the aging population. Nowadays life expectancy is higher than ever before in history, and keeps increasing. This, in turn, increases the socioeconomic burden of many neurological diseases. Many common molecular mechanisms affect different disorders through, for example, protein degradation, mRNA decay or gene regulation. Depending on the gene in question, its dysfunction may lead to a variety of conditions. Further, I will discuss molecular and genetic mechanisms particular to neurological diseases studied in this work.

1.2. Alzheimer's disease

1.2.1. Background

Alzheimer's disease (AD; OMIM #104300) is the most common neurodegenerative dementia in the elderly. According to the World Alzheimer's Report 2018, 50 million people are living with dementia worldwide. The total estimated cost of dementia around the world amounts to US\$1 trillion. This poses a significant burden due to an aging population.

AD was famously first described by a German psychiatrist and neuropathologist Alois Alzheimer in 1906 (Alzheimer, 1907). 70 years later it was recognised as a frequent cause of dementia and considered as one of the most common causes of death in the United States (Katzman, 1976). AD affects between 5-10 % of the population over 65 years of age (Rice *et al.*, 2001). Its prevalence increases from 1 % between 65 and 69 years of age up to 40-50 % in those older than 95 years (Nussbaum and Ellis, 2003). Some early symptoms include difficulty remembering names or recent events, apathy and depression. Later symptoms include disorientation, confusion, difficulty speaking, walking and swallowing (Alzheimer's Association, 2018).

Early onset AD (EOAD) was shown to be essentially identical to the most common form of senile dementia, the late onset AD (LOAD) (Terry and Davies, 1980). AD is considered early onset when it appears before the age of 65. The rate of EOAD in AD is generally considered to be between 1-2 %, however Zhu *et al.* (2015) showed it to be as high as 5.5 %, with the familial form accounting for 30 % of all EOAD cases. Both EOAD and LOAD are characterised by intracellular hyperphosphorylated tau protein aggregates called neurofibrillary tangles (NFTs) and extracellular senile plaques composed primarily of clumps of amyloid- β (A β) peptide (Grundke-Iqbal *et al.*, 1986; Sennvik *et al.*, 2000). Hardy and Higgins (1992) put forward the amyloid cascade hypothesis which claims that the deposition of A β is causing AD pathology and NFTs. Neurodegeneration, vascular damage and dementia are its result. Loss of neuropil, shrinkage of nerve fibre tracts, neurons or neuron death lead to atrophy of brain tissue starting at the entorhinal cortex, a region of the medial temporal lobe (Luxenberg *et al.*, 1987; Jobst *et al.*, 1992; Jobst *et al.*, 1994; Smith, 2002). It has since been speculated that this pathway may work in parallel with the aggregation of hyperphosphorylated tau protein, enhancing their toxic effects (Scheltens *et al.*, 2016). Kametani and Hasegawa (2018) suggest tau itself is the main factor in AD development and progression and not A β . At the same time another hypothesis claims the best target for therapeutic intervention in AD is the synaptic dysfunction affected by A β rather than the plaques themselves (Marsh and Alifragis, 2018).

AD is a multifactorial disease with a dichotomous pattern of inheritance. Approximately 70 % of the causes are genetic and the rest environmental (Dorszewska *et al.*, 2016). While some EOAD genetic causes are described as autosomal dominant, LOAD has no consistent Mendelian inheritance pattern (Bertram and Tanzi, 2005; Tanzi and Bertram, 2005). Mutations in three genes: *APP* (amyloid precursor protein), *PSEN1* (presenilin 1) and *PSEN2* (presenilin 2), are the most common causes of early onset familial AD (EO-fAD) and allele ϵ 4 of apolipoprotein E (*APOE*) is a well-established risk factor of LOAD. 30 to 50 % of inheritability of AD can be accounted for by changes in these four genes (Tanzi, 2012). Many other candidate genes were proposed through multiple studies such as genome-wide association studies (GWAS).

1.2.2. Epidemiology

Prevalence and incidence data can inform decision making in public health planning or diagnostics, and help compare the burden of disease at different locations or time periods. Incidence rates can be more useful for etiologic studies comparing varying rates

between different groups. Epidemiological studies need to include large numbers of data in order to be reliable sources of information, representing the general population.

Some AD risk factors include age (Guerreiro and Bras, 2015), female sex (Mielke *et al.*, 2018) and allele $\epsilon 4$ of *APOE* (Roses, 2006; Sando *et al.*, 2008). Among environmental risk factors are air pollution (Moulton and Yang, 2012; Peters *et al.*, 2019), toxic metals (Huat *et al.*, 2019), pesticides (Yan *et al.*, 2016), chronic psychological stress (Justice, 2018), brain trauma (Ramos-Cejudo *et al.*, 2018), hypothermia (Whittington *et al.*, 2010), starvation (Yanagisawa *et al.*, 1999), low level of education (Sharp and Gatz, 2011), obesity (Pegueroles *et al.*, 2018), diabetes mellitus (Zhang *et al.*, 2017) and smoking (Durazzo *et al.*, 2014).

1.2.2.1. Prevalence

Prevalence in epidemiology is the proportion in a given population or number of cases of a disease at a given point in time. 50 million people are estimated to currently live with dementia globally. In 2005, Ferri *et al.* reported a prevalence of dementia in North America equal to 6.4 % of the population at ≥ 60 years of age and 5.4 % in Western Europe for the same age group. That number was 4.6 % in Latin America and 4.0 % in China and developing Western Pacific. In 2001, 60.1 % of people with dementia lived in developing countries, and that number was predicted to rise to 71.2 % by the year 2040. A 100 % increase in prevalence rates was predicted to occur between 2001 and 2040 for developed countries, while a 300 % increase was predicted for China, India and developing Western Pacific.

Exponential growth can be observed for prevalence rates for AD after the age of 65. Between ages 60 and 85, almost a 15-fold increase in the prevalence of dementia is reported (Figure 1.1; Evans *et al.*, 1989).

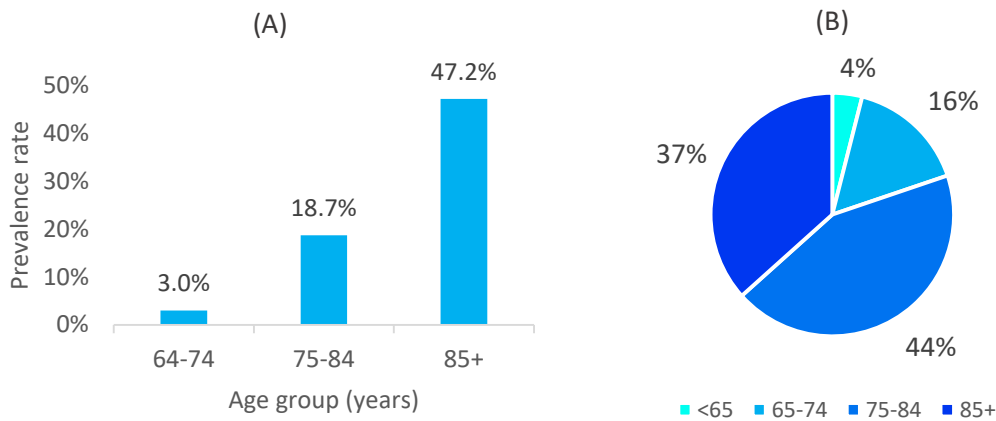


Figure 1.1. Prevalence rates of Alzheimer's disease for different age groups.

(A) Prevalence rates for probable AD (data from Evans *et al.*, 1989), (B) Proportion of AD dementia for different age groups in the United States, 2018 (percentages do not total 100 due to rounding; data from Alzheimer's Association, 2018).

An important risk factor for AD, which can be observed by looking at prevalence data, is sex. In a meta-analysis of prevalence of AD in Europe, it was shown that the proportion was 3.31 % for men and 7.13 % for women (Niu *et al.*, 2017). This may be partly due to the fact that the aged population is to a greater extent comprised of women than men. Yet, there have been reports suggesting a higher penetrance of certain mutations in women (Rao *et al.*, 1996; Cellini *et al.*, 2009).

Aging population, growing life expectancy and drastic prevalence increase of dementia after the age of 65 indicate a growing problem and an important socioeconomic burden.

1.2.2.2. Incidence

Incidence is the number of new cases of a condition per unit of time and, occasionally, for fraction of the population. It is a better measure for use in etiologic studies since it is not affected by differences in patient survival and its rate can be compared between different subgroups with different exposures (Ward, 2013). Although accurate incidence rates are difficult to establish due to requirement of large cohorts, long follow-up periods, determination of the age of onset and definition of a disease-free population, the studies show a consistent trend for AD patients. An incidence rate of 0.5 % per year for individuals aged 65-70 increases to 6-8 % for those older than 85 (Figure 1.2; Mayeux and Stern, 2012).

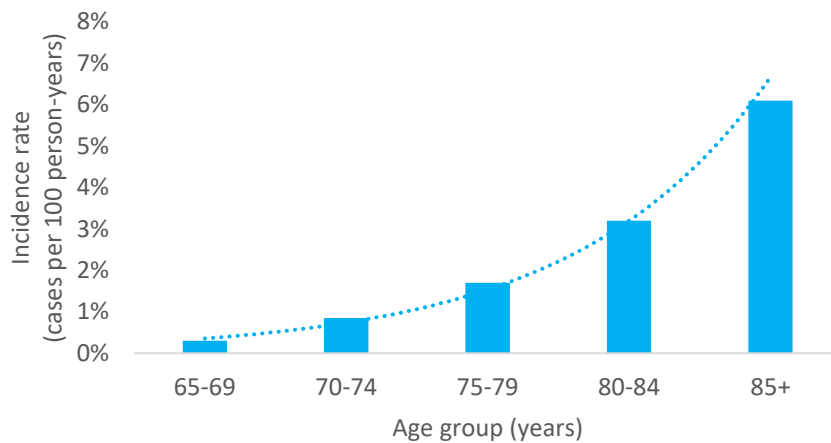


Figure 1.2. Annual incidence rate (per 100 person-years) of AD (data from Mayeux and Stern, 2012).

A trend line was added to show exponential growth.

High prevalence can be accounted for by extended duration of the disease and the rapid increase in incidence rate with age. The estimated annual incidence per 1000 person-years was 8.8 for Western Europe, 10.5 for North America, 9.2 for Latin America and 8.0 for China and developing Western Pacific (Ferri *et al.*, 2005). The meta-analysis of articles evaluating the European population points to an even higher incidence of 11.08 cases per 1000 person-years (8.97 for Southern European countries and 15.94 for Northern European countries). In this territory, women had a higher incidence of AD than men, 13.25 and 7.02 cases per 1000 person-years, respectively (Niu *et al.*, 2017). This is consistent with the prevalence data on difference between sexes.

1.2.2.3. Alzheimer's disease epidemiology in Spain

Studies of prevalence and incidence of AD in Spain point to a prevalence rate between 4.5-10.6 % and incidence of 5.4-10.8 per 1000 person-years (Manubens *et al.*, 1995; López-Pousa *et al.*, 2004; Bermejo-Pareja *et al.*, 2008; Lobo *et al.*, 2011; Tola-Arribas *et al.*, 2013). In a cross-sectional study by Tola-Arribas *et al.* (2013), which was comprised of screening and diagnostic confirmation phases and performed in a door-to-door manner in Valladolid, Spain, the prevalence of AD was 2.2 % for men and 5.8 % for women. People from the study aged 65-69 had a 0.7 % prevalence of AD with that number raising to 9.9 % for those aged 80-84.

In a study by Lobo *et al.* (2011) individuals aged 55 and older, living in Zaragoza, Spain, were studied. Rapid increase of incidence was observed with age (Figure 1.3).

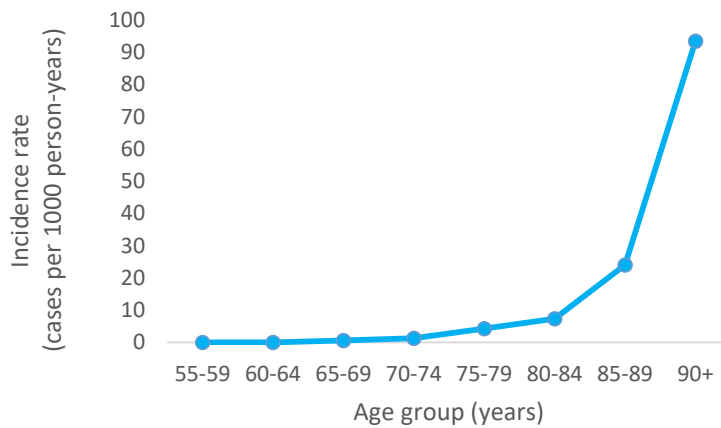


Figure 1.3. Incidence rate (per 1000 person-years) of Alzheimer’s disease in a Spanish cohort from Zaragoza (data from Lobo *et al.*, 2011).

Again, incidence rates were significantly higher for women than men (Figure 1.4), with 6.8 cases per 1000 person-years for the former and 3.7 for the latter.

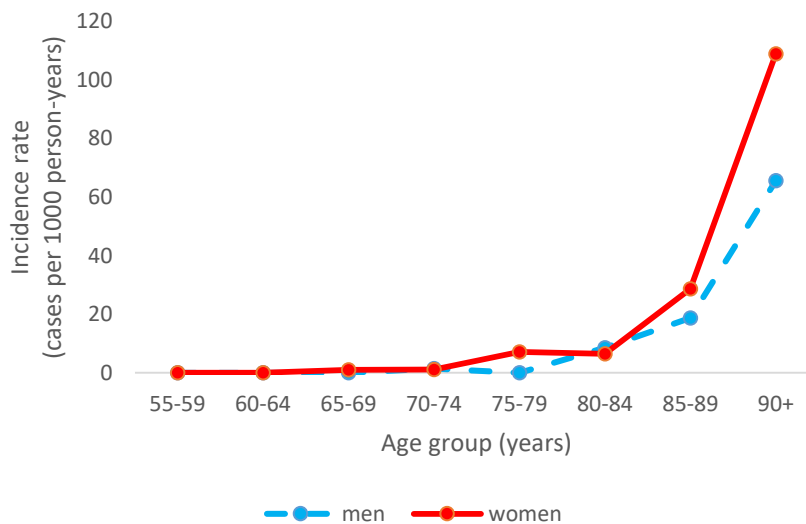


Figure 1.4. Comparison of incidence rates (per 1000 person-years) of Alzheimer’s disease between sexes of a Spanish cohort from Zaragoza (data from Lobo *et al.*, 2011).

1.2.3. Clinical features and treatment

The symptoms of AD can be divided into cognitive: memory loss, apraxia, etc. and non-cognitive: mood swings, delusions, etc. (Awada, 2015). National Institute on Aging and the Alzheimer's Association proposed revised guidelines for AD diagnosis in 2011 (Albert *et al.*, 2011; Jack *et al.*, 2011; McKhann *et al.*, 2011). These revised guidelines include the previous 1984 criteria based on assessment of symptoms, cognitive test results, neurological evaluation and reports from the individual and the relatives but differ by incorporation of biomarker tests (cerebrospinal fluid (CSF) analysis, functional and structural imaging and amyloid positron emission tomography) and formalisation of different stages of the disease.

Short term memory deficits, difficulty with language and finding words are primary clinical features (Becker *et al.*, 1994). Among non-cognitive symptoms, mood dysfunction, depression, apathy, anxiety, insomnia, and agitation or psychosis are prevalent. AD dementia is the clinical manifestation of the pathophysiological process of AD. For a positive diagnosis of dementia, the person presenting the symptoms needs to show functional impairment and behavioural or cognitive impairment that is represented in at least two of the domains: memory, visuospatial abilities, language, personality, behaviour or executive function. For a diagnosis of probable AD dementia, the above dementia criteria have to be met, the onset of the disease needs to be insidious and progressive, cognitive impairment and amnesic or nonamnesic symptoms need to be reported (Ulep *et al.*, 2017).

EOAD occurs before the age of 65 years and is caused primarily by mutations in three genes: *APP*, *PSEN1* and *PSEN2*. The main genetic risk factor of LOAD is allele $\epsilon 4$ of *APOE*. While the neuropathology features of both sporadic and monogenic AD are similar, phenotype of monogenic cases of EOAD is broader with disturbances in behaviour, myoclonus, epilepsy and focal presentations. Survival in monogenic or sporadic AD is 6-9 years from diagnosis, on average (Pilotto *et al.*, 2013).

The only classes of drugs approved by the US Food and Drug Administration (FDA) for treatment of AD symptoms are cholinesterase inhibitors (ChEI) and low-affinity *N*-methyl-D-aspartate receptor (NMDA) antagonists. Loss of cholinergic neurons, and thus cholinergic neurotransmission, contributes to AD symptoms. Donepezil, rivastigmine and galantamine, all ChEI, enhance cholinergic neurotransmission. The NMDA receptor antagonist, memantine, reduces glutamatergic neuronal excitotoxicity (Ulep *et al.*, 2017). Currently there is no treatment available that could cure or arrest the progression of the disease.

1.2.4. Alzheimer's disease pathology

AD neuropathological hallmarks are intracellular NFTs and extracellular amyloid (senile) plaques. Other core features involve synaptic deterioration and neuronal death (Ballard *et al.*, 2011). The amyloid cascade hypothesis recons the deposition of A β peptide is the cause of AD pathology resulting in formation of NFTs, neurodegeneration, vascular damage and dementia (Hardy and Higgins, 1992).

1.2.4.1. Amyloid plaques

Amyloid plaques (Figure 1.5) are mainly composed of aggregates of A β peptide, a 36-43 amino acid protein. A β is formed after proteolytic cleavage of APP, a transmembrane protein. The cleavage may occur through two different pathways. In the first pathway, α -secretase cleaves APP at the extracellular N-terminal to form a soluble form of APP known as sAPP α . The remaining fragment of APP, still attached to the membrane, can further be cleaved by γ -secretase to form a soluble p3 fragment and a short polypeptide, APP-intracellular domain (AICD), still embedded in the membrane. The second pathway involves cleavage of APP by β -secretase forming sAPP β . Further cleavage of the remaining membrane-associated fragment by γ -secretase produces A β peptides (Figure 1.6; Chasseigneaux and Allinquant, 2012). The main products of this cleavage are A β ₄₀ and A β ₄₂ isoforms of the peptide. According to Palop and Mucke (2010) expression of A β ₄₂ isoform is increased by mutations correlated with familial AD and the overexpression of A β decreases synaptic transmission.

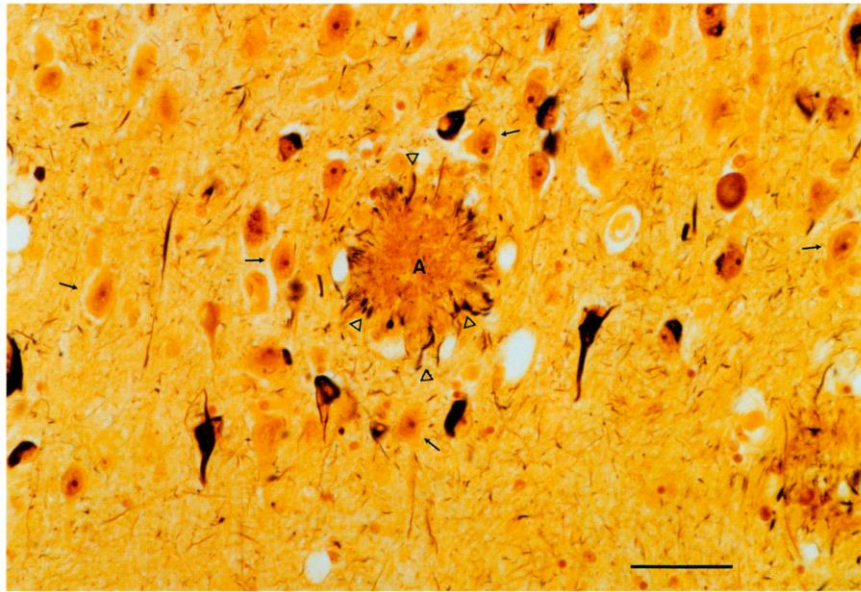


Figure 1.5. Hallmark lesions of Alzheimer's disease visualised by modified Bielschowsky Silver Stain (from Selkoe *et al.*, 1991).

In the center, denoted with an A, a large senile plaque with structurally abnormal neurites (open arrowheads). Open arrowheads point to structurally abnormal (dystrophic) neurites, the arrows point to cytologically normal neurons while dark neurofibrillary tangles are visible in the cytoplasm of pyramidal neurons. Bar = 50 μ m. See appendix I for license agreement.

1.2.4.2. Neurofibrillary tangles

NFTs are composed of aggregated tau protein. In normally functioning neurons, tau is a microtubule-associated protein found primarily in axons. It stabilises microtubules and helps with their assembly by interaction with tubulin (Binder *et al.*, 1985; Medeiros *et al.*, 2011). Microtubules disassemble due to hyperphosphorylation of tau which causes its fibrillation and the aggregation of fibrils to form NFTs (Figure 1.6; Alonso *et al.*, 2018). This process takes place in AD affected individuals although it is not clear whether it is the cause of the disease or its manifestation. The helical filamentous NFTs are mainly deposited in the neurons of the medial temporal lobe, lateral parietotemporal region and frontal association cortices (Ulep *et al.*, 2017). In contrast with the senile plaques, the location and density of NFTs has been proven to be correlated with progression of AD in patients (Ballard *et al.*, 2011). Although the amyloid cascade hypothesis suggests it is the A β aggregates that influence formation of NFTs, it has since been speculated that these pathways may also work in parallel, enhancing their toxic effects (Scheltens *et al.*, 2016).

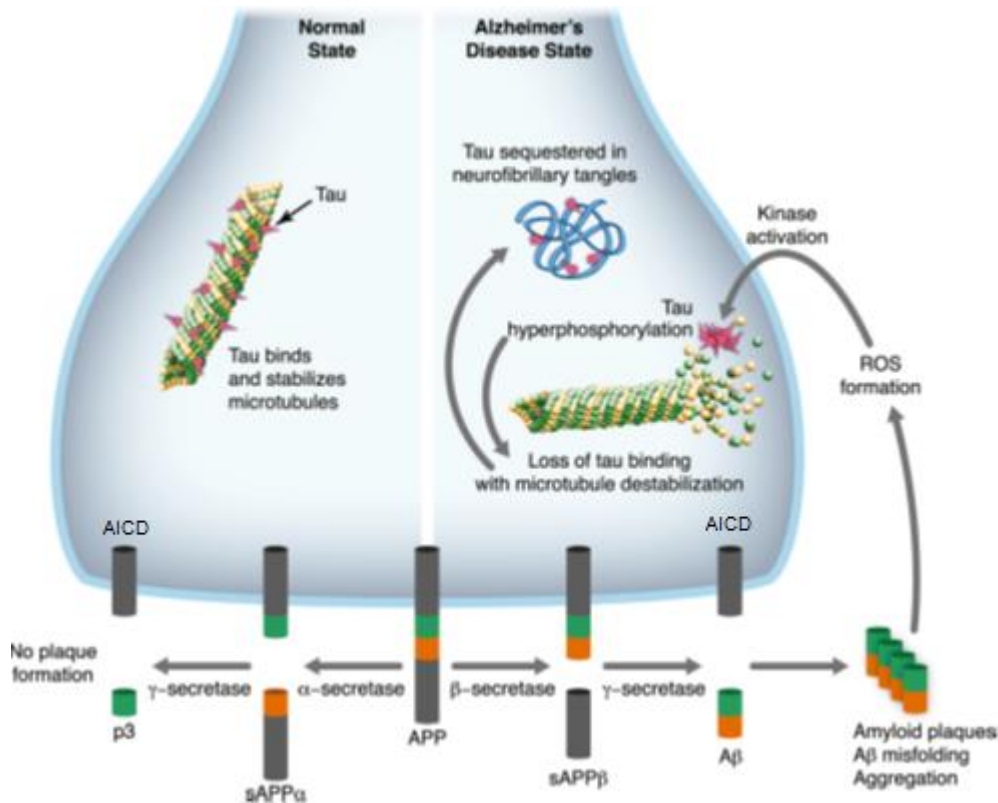


Figure 1.6. Comparison of normal state and Alzheimer's disease state, with amyloid β aggregates and neurofibrillary tangles formation as the main hallmarks of the disease (from Gireud *et al.*, 2014).

See appendix I for license agreement.

1.2.4.3. Progression of Alzheimer's disease pathology through the brain

Scahill *et al.* (2002) mapped the evolution of regional atrophy in AD by use of volumetric magnetic resonance imaging (MRI). Hippocampal atrophy was observed even in presymptomatic subjects and mildly affected patients. The medial parietal lobe atrophy could be observed at all stages of the disease while the frontal lobe atrophy was observed in more advanced stages. An important observation was that regional atrophy occurs even before the clinical onset of symptoms and spreads from the medial temporal lobe to other parts of the brain in a specific manner. Of note is also that the preclinical stage of AD precedes the clinical symptoms potentially by many years, even decades (Dubois *et al.*, 2016). Entorhinal cortex of the parahippocampal gyrus, which was suggested as the first site of expression of AD pathology by Braak and Braak (1991), shows gravest atrophy. The sequence of pathology of AD spreading, together with the timeline proposed by Ohm *et al.* (1995) are summarized in figure 1.7.

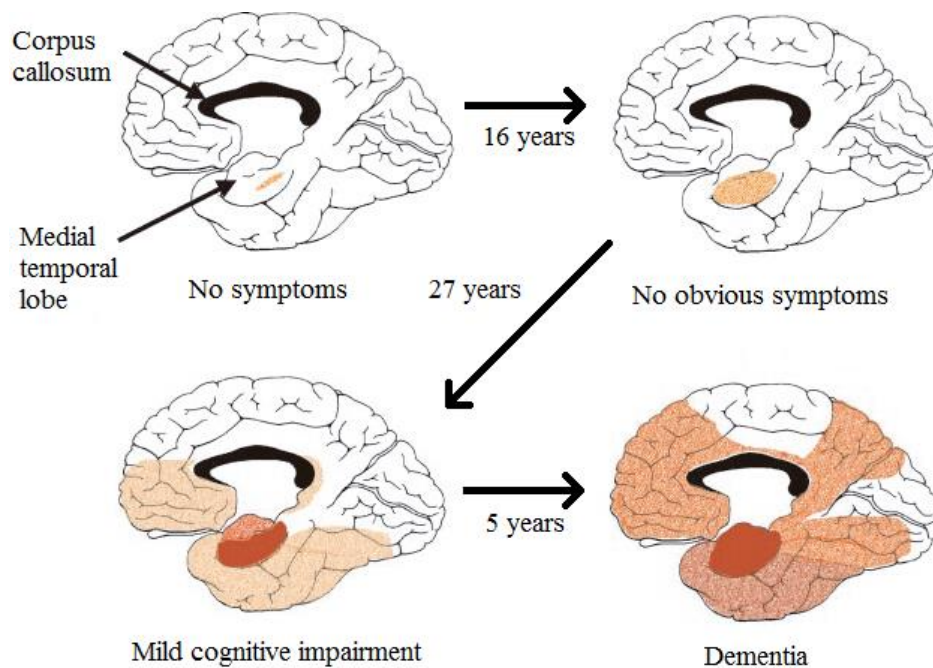


Figure 1.7. Spread of neurofibrillary pathology in Alzheimer’s disease (modified from Smith, 2002).

Density of tangles is proportional to depth of colour. See appendix I for license agreement. Copyright (2002) National Academy of Sciences, USA.

The atrophy of an AD affected brain was shown to be significant when comparing the *cornu Ammonis* of a hippocampus of a non-affected elderly person and late stage AD patient. The former occupied 1.5 ml and contained about 9 million neurons and the latter about 0.5 ml in volume and some 1.44 million neurons (Bobinski *et al.*, 1996).

1.2.5. Etiology

1.2.5.1. Environmental factors

Many environmental factors have been proposed as possible risk factors for LOAD to date. Some of them include air pollution (Moulton and Yang, 2012; Peters *et al.*, 2019), toxic metals (Huat *et al.*, 2019), pesticides (Yan *et al.*, 2016), chronic psychological stress (Justice, 2018), brain trauma (Ramos-Cejudo *et al.*, 2018), hypothermia (Whittington *et al.*, 2010), starvation (Yanagisawa *et al.*, 1999), low level of education (Sharp and Gatz, 2011), obesity (Pegueroles *et al.*, 2018), diabetes mellitus (Zhang *et al.*, 2017) and smoking (Durazzo *et al.*, 2014).

Environmental factors, which account for about 30 % of AD cases (Dorszewska *et al.*, 2016), cause stress on systemic, inter and intracellular levels. These may cause neuron or synapse loss and cognitive impairment through many different pathways (Wainaina *et al.*, 2014).

1.2.5.2. Genetic basis of Alzheimer's disease

Genetic causes account for approximately 70 % of all AD cases. EO-fAD is primarily caused by mutations in *APP*, *PSEN1* and *PSEN2* genes. Little is known about the function of APP, yet it is speculated it is involved in neuron migration during early development in the brain. Cleavage of APP by α -, β - and γ -secretases leads to formation of peptides with further functions. sAPP may be involved in neuron formation and growth as well as acting as an inhibitor for certain proteins, while A β plays a role in synaptic plasticity (Caillé *et al.*, 2004; Parihar and Brewer, 2010). PSEN1 is the catalytic subunit of the γ -secretase complex and PSEN2 is its homolog. They are thus closely involved with APP cleavage and formation of A β (De Strooper, 2003).

According to Cruts *et al.* (2012) the recorded number of mutations for the three genes involved in a fully penetrant, autosomal dominant mode of inheritance was: 32 mutations (24 plus duplications) for *APP*, 185 for *PSEN1* and 12 for *PSEN2* with an additional not fully penetrant mutation (*PSEN2*^{N141I}). Most of these mutations lead to an increase in the ratio of A β ₄₂ to A β ₄₀ and subsequently aggregation of amyloid fibrils (Jarrett *et al.*, 1993; Tanzi and Bertram, 2005). Other candidate genes for EO-fAD involve *MAPT* encoding tau protein, PAX transcription activation domain interacting protein gene (*PAXIP1*) and *PSENEN* encoding Presenilin Enhancer Protein 2, another γ -secretase component (Tanzi, 2012).

A well-established genetic risk factor for LOAD is allele $\epsilon 4$ of *APOE* (Strittmatter *et al.*, 1993). *APOE* protein forms lipoproteins which function as cholesterol and other fats transporters in the blood. Except cholesterol metabolism and lipid transport, *APOE* is also involved in inflammation control, synaptic function, neurogenesis and the generation and trafficking of $A\beta$ and APP (Giri *et al.*, 2016). Disruption of normal cholesterol levels may cause disorders affecting the cardiovascular system. There are three major *APOE* alleles: $\epsilon 2$, $\epsilon 3$, $\epsilon 4$, with the $\epsilon 3$ being the most common, the $\epsilon 4$ second and $\epsilon 2$ the least common (Table 1).

Table 1.1. Frequency of *APOE* genotypes in a US population.

Genotype	Population (%)
$\epsilon 2/2$	1
$\epsilon 2/3$	12
$\epsilon 3/3$	60
$\epsilon 3/4$	21
$\epsilon 4/4$	2

Subjects with genotype $\epsilon 2/\epsilon 4$ are not included due to its low frequency. Table uses an estimate of 46 million people from US over 60 y/o from year 2000 (data from Raber *et al.*, 2004).

The *APOE* alleles result from amino acid combinations at residues 112 and 158. Allele $\epsilon 2$ has the following combination: Cys₁₁₂/Cys₁₅₈, allele $\epsilon 3$: Cys₁₁₂/Arg₁₅₈ and allele $\epsilon 4$: Arg₁₁₂/Arg₁₅₈. While allele $\epsilon 4$ increases the AD risk 4-fold if inherited in one copy and more than 10-fold if in two, allele $\epsilon 2$ is considered to encode a neuroprotective variant of the protein (Tanzi, 2012). Increasing number of *APOE* $\epsilon 4$ alleles is associated with an earlier onset of the disease. AD patients who were *APOE* $\epsilon 4$ homozygotes had a lower volume of grey matter with age and higher hippocampal atrophy, as well as increased amyloid load and cerebral amyloid angiopathy (Giri *et al.*, 2016). Higher number *APOE* $\epsilon 4$ alleles is also associated with increased levels of phosphorylated tau protein (Han *et al.*, 2010) and higher $A\beta$ accumulation (Lim and Mormino, 2017). On the other hand, carriers of *APOE* $\epsilon 2$ had a lower concentration of tau in the CSF and lesser rate of hippocampal

atrophy. It is speculated that APOE ϵ 2 is more efficient at clearing A β than APOE ϵ 4 due to its higher affinity for the peptide (Giri *et al.*, 2016).

Many other genes were identified as possible risk factors for AD through studies made by genetic linkage analysis, study of candidate genes, GWASs and next-generation sequencing (NGS) technologies. The identified genes can be grouped into clusters involving three major pathways: Inflammatory response, endocytosis and lipid metabolism (Table 2).

Table 1.2. Selected genes and their variants associated with Alzheimer’s disease.

Pathway	Gene	Chromosome	SNP	Function
Axonal transport and cytoskeletal function	<i>NME8</i>	7p14.1	rs2718058	Ciliary function Neuronal cell proliferation
Cholesterol metabolism and immune response	<i>CLU</i>	8p21-p12	rs11136000	Synapse turnover Complement regulation Chaperone protein
Cytoskeletal function and axonal transport	<i>CASS4</i>	20q13.31	rs7274581	Cell migration Cell adhesion
			rs6024870 rs16979934	
Endocytosis Lipid metabolism	<i>CELF1</i>	11p11	rs10838725	mRNA editing Pre-mRNA splicing
			<i>SORL1</i>	11q23.2-q24.2
Endocytosis Synapse function	<i>CD2AP</i>	6p12	rs9296559 rs9349407	Receptor-mediated endocytosis Cytokinesis Cytoskeletal dynamics
			<i>PICALM</i>	11q14
Epigenetic regulation	<i>ZCWPW1</i>	7q22.1	rs1476679	Epigenetic regulation Neural development

Table 1.2. Selected genes and their variants associated with Alzheimer’s disease (continued).

Hippocampal synaptic function	<i>PTK2B</i>	8p21.1	rs28834970	Calcium homeostasis MAP kinase signaling
	<i>CRI</i>	1q32	rs3818361 rs6656401	Amyloid β clearance Complement activation
	<i>CD33</i>	19q13.3	rs3865444 rs3826656	Clathrin-mediated endocytosis Cell signaling
	<i>MS4A</i>	11q12.2	rs610932 rs670139 rs4938933	Signal transduction Immune function
Immune response	<i>TREM2</i>	6p21.1	rs75932628	Inflammatory response
	<i>EPHA1</i>	7q34	rs11771145 rs11767557	Synaptic development Immune function Neural development
	<i>HLA-DRB5/HLA-DRB1</i>	6p21.3	rs9271192	Immune function Histocompatibility
	<i>INPP5D</i>	2q37.1	rs35349669	Cytokine signaling Immune function
Immune response Hippocampal synaptic function	<i>MEF2C</i>	5q14.3	rs190982	Myogenesis Synapse formation
Immune response Lipid metabolism	<i>ABCA7</i>	9p13.3	rs3764650 rs3752246	Phagocytosis Lipid homeostasis
	<i>APOE</i>	19q13.2		Lipid transport Synaptic vesicle endocytosis Cytoskeletal dynamics
	<i>PLD3</i>	19q13.2	rs145999145	Signal transduction Epigenetic modification

Table 1.2. Selected genes and their variants associated with Alzheimer’s disease (continued).

Signal transduction	<i>UNC5C</i>	4q22.3	rs137875858	Neural development
Synaptic vesicle endocytosis	<i>BINI</i>	2q14.3	rs744373 rs7561528	Synaptic vesicle endocytosis APP trafficking Cytoskeletal dynamics
Synapse function endocytosis	<i>SLC24A4/RIN3</i>	14q32.12	rs10498633	Cell signaling Neural development
Tau pathology Angiogenesis	<i>FERMT2</i>	14q22.1	rs17125944	Cell–cell adhesion Angiogenesis
	<i>DSG2</i>	18q12.1	rs8093731	Cell–cell adhesion
	<i>AKAP9</i>	7q21-q22	rs144662445 rs149979685	
	<i>ADAM10</i>	15q22	rs2305421	Hippocampal neurogenesis Cell adhesion

SNP (single nucleotide polymorphism; data from Giri *et al.*, 2016).

Together with *APOE*, genes *CLU*, *ABCA7* and *SORL1* have been indicated as AD risk factors involved in cholesterol metabolism. As shown by Solomon *et al.* (2009), high cholesterol levels in midlife increased the risk of AD and vascular dementia in older age. Among genes associated with an increased risk of AD and involved in an immune response are *CRI*, *CD33*, *MS4A*, *EPHA1*, *TREM2*, *HLA-DRB5/HLA-DRB1*, *INPP5D* and also *CLU* and *ABCA7*. The immune system provoking neuroinflammation, a hallmark of AD, is involved in LOAD pathogenesis as shown by several GWASs (Bertram *et al.*, 2008; Harold *et al.*, 2009). Endocytosis related genes identified by GWASs as probable risk factors for AD are, among others: *BINI*, *CD2AP*, *PICALM*, *EPHA1*, *PTK2B* and *SORL1*. Not only are some of these directly involved in trafficking APP but endocytosis itself is important for response to neural damage and synaptic transmission (Giri *et al.*, 2016). Other genes proposed as risk factors for AD, and not directly involved in the aforementioned pathways, are *MEF2C* which plays a role in myogenesis (Dodou *et al.*, 2003); *CASS4*, *NME8* and *CELF1* which are involved in axonal transport and function of the cytoskeleton; *FERMT2*

implicated in angiogenesis and tau pathology (Ghani *et al.*, 2018); *ZCWPW1* an epigenetic regulator, *SLC24H4-RIN3* taking part in cell signalling; and *DSG2* with a role in cell to cell adhesion (Giri *et al.*, 2016). Some rare variants associated with a risk of AD have been identified using whole-exome or whole-genome sequencing (WES and WGS respectively). These include *TREM2*, *PLD3*, *UNC5C*, *AKAP9*, and *ADAM10*.

As previously mentioned the complexity of AD and its multifactorial nature makes it difficult to treat. There are many pathways to follow due to strong and diverse genetic component involvement in its etiology. Yet three main pathways seem to encapsulate most of the genetic risk factors: the immune response, lipid metabolism and endocytosis. Next, and due to its relevance for this thesis, I will give more details on the complement system.

1.2.6. Complement system

An important part of the immune response related to AD is found in the complement system. It consists of a number of plasma and cell membrane proteins that induce an inflammatory response. It is an effector of the humoral immunity implied in cytolysis. Functions of the complement include recognition and killing of pathogens. The complement is phylogenetically conserved among both invertebrates and mammals suggesting its high importance (van Beek *et al.*, 2003). There are three different routes of activation of the complement system, the classical pathway, the alternative pathway and the mannose-binding lectin (MBL) pathway (Figure 1.8).

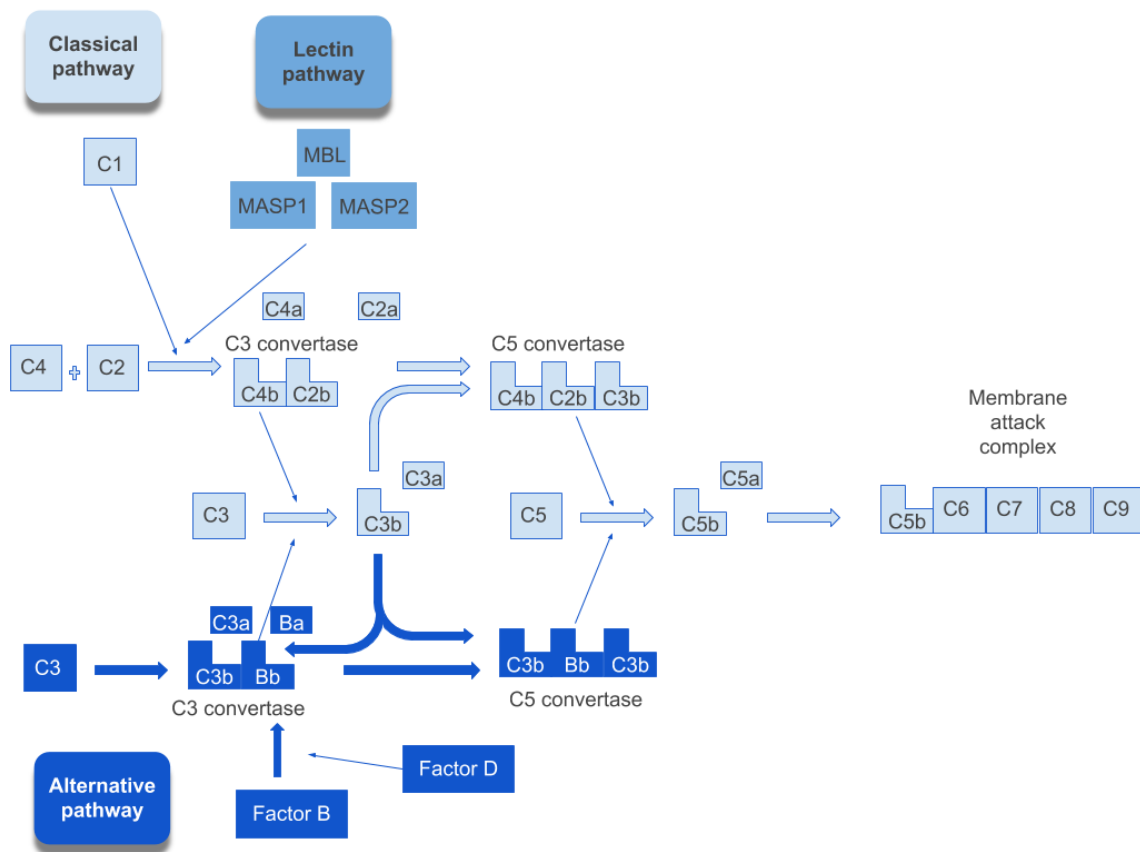


Figure 1.8. Pathways of the complement cascade.

Detailed explanation of the figure in text.

Activation of the classical pathway involves initiation, which is done by binding of antigen-antibody complex to C1q pattern recognition molecule of the C1 complex. C1 is composed of the C1q molecule and a heterotetramer C1r₂s₂ made up of two C1r and two C1s proteases. C1q binding of the activation molecule provokes a conformational change and autoactivation of C1r protease which in turn activates C1s. Activated C1s then cleaves C2 and C4 to C2a, C2b and C4a, C4b respectively. C4bC2b forms a complex called C3 convertase that cleaves C3 to C3a and C3b. C3a proinflammatory anaphylatoxin recruits leukocytes by interacting with a G-protein–coupled receptor C3aR. C3b forms a complex with C4bC2b (C4bC2bC3b) called C5 convertase. C5 is cleaved by this convertase to C5a and C5b. Just like C3a, the C5a is a proinflammatory anaphylatoxin interacting with C5aR receptor and attracting leukocytes. C5b together with C6, C7, C8 and C9 form the membrane attack complex (MAC) which incorporates in the cell membranes of the opsonized by C3b target cells and provokes their lysis by pore formation (Noris and Remuzzi, 2013; Mortensen *et al.*, 2017).

Although the alternative pathway is activated by spontaneous hydrolysis of C3, it is controlled by the inhibitor Factors H and I, since they do not inactivate surface bound C3b but do so with plasma C3b. Hydrolysed C3 (C3b) binds to the serine protease Bb, formed by cleavage of circulating Factor B into Ba and Bb by Factor D. C3bBb complex acts as a C3 convertase cleaving C3 into C3a and C3b. More C3bBb complexes are then formed thus initiating an amplification process as these act as C3 convertases in the alternative pathway. C3b deposits on the surface of the targeted cells opsonizing them. It can then bind properdin, which protects it from being inactivated by Factors H and I. These factors inhibit the alternative pathway inactivating C3b in solution. C3bBb complex can bind another C3b molecule forming a C5 convertase. This in turn takes part in MAC production and lysis pathway. The constant spontaneous activation of C3 allows for a fast reaction of the alternative pathway without the necessity of antibody involvement (Noris and Remuzzi, 2013).

Finally, the MBL pathway only differs from the classical pathway by the components involved in its initiation. Mannose-binding lectin binds mannose residues on pathogen surface and the associated serine proteases MASP1 and MASP2 are activated. C2 and C4 are then cleaved by these proteases to form C2a, C2b and C4a, C4b respectively. The following steps of this pathway are identical to those of the classical one. Ficolins can also initiate this pathway. Ficolins bind carbohydrate moieties on pathogen surface and are also associated with MASPs (Endo *et al.*, 2011; Noris and Remuzzi, 2013).

1.2.6.1. Complement system in Alzheimer's disease pathogenesis

The endothelial cells of microvessels, pericytes and astrocytes forming the blood-brain barrier (BBB) separate the brain parenchyma from proteins found in blood plasma making it an immune privileged site. As most of the complement proteins are produced in the liver it was first assumed only a break in the BBB could allow for these proteins to affect the brain. Later Levi-Strauss and Mallat (1987) discovered murine astrocytes are capable of producing components of the alternative pathway. Now we know microglia, neurons and oligodendrocytes can also produce complement proteins (Barnum, 1995; Gasque *et al.*, 2000). While astrocytes and microglia express complement inhibitors, neurons seem susceptible to attack by this pathway (Singhrao *et al.*, 2000). Complement mRNA has been shown to be upregulated in AD brain, with C1q mRNA being 11- to 80-fold higher than in controls (Yasojima *et al.*, 1999). Complement factors were also identified in amyloid plaques (Eikelenboom and Stam, 1982) and it is postulated that A β

binds C1q and activates the classical complement pathway *in vivo*. It was also shown that purified NFTs activate complement system in plasma, making aggregated tau a possible antibody-independent activator of this pathway (Shen *et al.*, 2001). Complement components can be found in the amyloid plaques from the earliest stages of A β deposition and their activation corresponds with clinical expression of AD. They can be involved in AD neuropathology by improper activation or inhibition of the complement pathways.

Neuroimmune regulatory proteins such as glycosylphosphatidylinositol (GPI) anchored molecules (CD24, CD90, CD55 and CD59), molecules of the immunoglobulin superfamily (CD22, Siglec-10, CD200, ICAM-5) and others (CD47, fractalkine, TAM receptor tyrosine kinase, C3a and Factor H) control negative immune responses (Shen *et al.*, 2013). Although Singhrao *et al.* (2000) reported the susceptibility of neurons to the complement pathway, human serum, containing complement proteins, was not toxic against human neurons. Homologous restriction, the presence of species-selective complement inhibitors on cell surface, is the proposed cause of this lack of susceptibility. This was confirmed by removing CD55 and CD59 from the SH-SY5Y neuroblastoma cells by treatment with phosphatidylinositol-specific phospholipase C (PIPLC) and further incubation with human serum containing complement (Shen *et al.*, 2013).

1.2.6.2. Complement receptor 1

A GWAS study involving almost 6 thousand AD patients and close to 8.5 thousand controls has identified complement receptor 1 (CR1; CD35) as a possible risk factor for LOAD (Lambert *et al.*, 2009). CR1 is a receptor for C3b and C4b molecules. It is a type-I transmembrane glycoprotein that has multiple functions in immune and glial-mediated inflammation. It regulates complement activation by interacting with C3b/C4b and is involved in phagocytosis of opsonized complexes. In addition to the membrane bound form, CR1 can also be soluble (sCR1). The transmembrane form has been found to be expressed on plasma membrane of erythrocytes, macrophages, monocytes, eosinophils, B-lymphocytes, subpopulation of CD4⁺ T cells, Langerhan cells, glomerular podocytes and dendritic cells. The soluble form was found in plasma and urine (Khera and Das, 2009).

CR1 has a long extracellular chain composed of 30 short consensus repeats (SCRs), each approximately 60 amino acids long, followed by a transmembrane region and a short cytoplasmic tail. The SCRs are divided into four long homologous repeats (LHRs) each consisting of 7 SCRs. Two SCRs closest to the transmembrane region do not cluster into an LHR. Each LHR is a functional domain. LHR-A primarily binds C4b, LHR-B and LHR-

C bind C3b/C4b, PfEMP1 (*Plasmodium falciparum* erythrocyte membrane protein 1 which is present on the membrane of erythrocytes infected by *Plasmodium falciparum*) and act as cofactors for Factor I. LHR-D binds MBL and C1q. The sCR1 has a similar structure, lacking the cytoplasmic tail as it is the product of cleavage of the transmembrane form of CR1. There are four isoforms of CR1 (CR1*1, CR1*2, CR1*3 and CR1*4), each with a different number of LHRs (Figure 1.9). Unequal crossing over of chromosomes causes insertions/deletions which are responsible for isoform formation (Holers *et al.*, 1987). The most common isoform, CR1*1 has 4 LHRs, second most common, CR1*2 has 5 with the LHR-B repeated. Other isoforms are very rare with CR1*3 having only three LHRs, lacking LHR-B and CR1*4 having 2 additional LHR-B (Khera and Das, 2009).

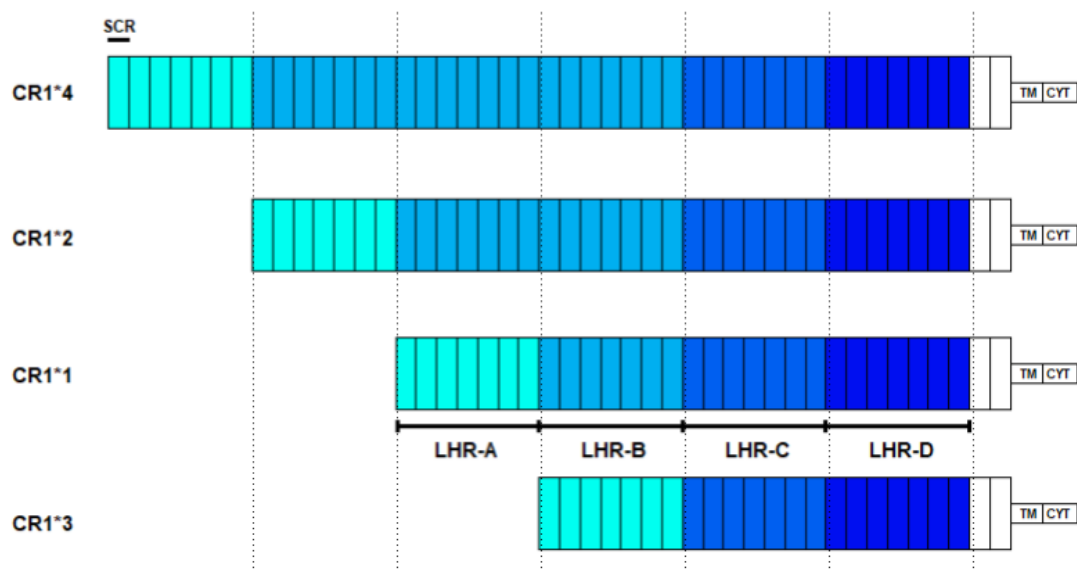


Figure 1.9. Schematic representation of CR1 isoform structures.

SCR – short consensus repeat; LHR – long homologous repeat; TM – transmembrane domain; CYT – cytoplasmic tail. Figure modified with permission from Kisserli *et al.*, 2017 licensed under CC BY-NC-ND 3.0.

CR1 functions as a regulator of the complement system by enhancing the decay of C3 and C5 convertases (Iida and Nussenzweig, 1981) and by being a cofactor to Factor I in the inactivation of C3b and C4b (Khera and Das, 2009). CR1 also plays a role in immune complexes clearance. Once such complex is opsonized by C3b or C4b, the erythrocyte-bound CR1 can attach to these molecules and carry them to the liver and spleen where they are transferred to macrophages and eventually metabolized. Immune complexes opsonized by C3b can also be detected by monocytic CR1. Together with Fc-gamma receptors, CR1

is responsible for uptake of these particles, which are later broken down in lysosomes. CR1 on B-lymphocytes is involved in the cells' proliferation. It is clear that CR1 is an important receptor for the immune response and the complement system regulation (Khera and Das, 2009).

The isoform found to be associated with AD risk is CR1*2. It has been shown that individuals with *CR1*1/CR1*2* genotype were 1.8 times more likely to be at risk of developing AD than those with *CR1*1/CR1*1* genotype. The longer isoform (CR1*2) was found to be expressed at lower protein levels than CR1*1. These findings suggest that CR1*2 may be connected to increased complement activation and lower A β clearance. In neurons, the pattern of expression of the CR1*1 isoform in *CR1*1/CR1*1* homozygotes is filiform while it shows vesicular-like profiles for *CR1*1/CR1*2* heterozygotes. Also, in neurons, CR1 is found in endoplasmic reticulum intermediate compartments in *CR1*1/CR1*1* genotype whereas it is found in lysosomes for *CR1*1/CR1*2* genotype. The processing of the two isoforms may thus be distinct (Hazrati *et al.*, 2012; Mahmoudi *et al.*, 2015).

Bin *et al.* (2019) found the expression of *CR1* to be significantly elevated in the entorhinal cortex and hippocampus of AD affected brains in comparison to controls. Both these regions are important in development of the disease.

Nine *CR1* single nucleotide polymorphisms (SNPs) are associated with AD risk (rs1408077, rs3818361, rs4844610, rs6656401, rs6691117, rs6701710, rs6701713, rs11803956 and rs116806486). CR1 is involved in clearance of A β , regulation of the complement cascade and its mRNA levels correlate with NFTs density and phosphorylated tau abundance (Zhu *et al.*, 2015). Alteration of A β clearance may affect amyloid plaques formation, the neuroinflammation caused by complement cascade plays an important role in neurodegeneration and *CR1* mRNA level correlation with NFTs density may indicate a relation between CR1 and tauopathy. It is proposed that through these mechanisms *CR1* and its SNPs are involved in AD risk.

1.3. Amyotrophic lateral sclerosis and primary lateral sclerosis

1.3.1. Amyotrophic lateral sclerosis background

Amyotrophic lateral sclerosis (ALS; OMIM #105400) is the most common motor neuron disease (MND) with a prevalence estimated at 222,801 cases worldwide (Arthur *et al.*, 2016). Although ALS is a rare disorder it carries a high socioeconomic burden. It is fatal, with an average survival time of 3 years and only about 5 % of patients living 20 years or longer. Since a clear cause of the disease is unknown, no effective cure has been developed.

First described by a French neurologist, Jean-Martin Charcot, in 1869 (Charcot and Joffroy, 1869; Goetz, 2000) it has become well known due to famous sufferers of the disease such as Lou Gehrig or Stephen Hawking. Characterized by degeneration of upper and lower motor neurons (UMNs and LMNs respectively) at spinal and bulbar level, it causes limb paralysis, dysarthria, dysphagia and respiratory failure. ALS is most commonly sporadic (90-95 %), although there is a familial ALS (fALS) form with primarily autosomal dominant inheritance. Typically, the onset of first symptoms is at the age of 50 to 65 years. Muscle weakness or twitching can be some of the early symptoms of the disease. There exist different phenotypical expressions of ALS with distinct onset patterns for UMNs and LMNs involvement: limb-onset ALS, bulbar onset ALS, and primary lateral sclerosis (PLS) and progressive muscular atrophy (PMA), both often considered as points in a continuum of ALS (Zarei *et al.*, 2015).

Behavioural and cognitive deficits may occur with ALS, ranging from mild, moderate to frontotemporal dementia (FTD; Strong *et al.*, 2003). The frontal lobe degeneration associated with ALS which causes dementia, suggests a continuum of ALS with cognitive or behavioural impairment and FTD. Some of the symptoms associated with FTD are changes in social behaviour and personality as well as impairments in language and executive function (Zago *et al.*, 2011).

Mutations in superoxide dismutase 1 (*SOD1*) gene, encoding an antioxidant enzyme, and in the protein encoded by the chromosome 9 open reading frame 72 (*C9orf72*) gene are the most common causes of ALS. Aggregation of misfolded SOD1 in motor neurons is one of the proposed mechanisms behind ALS. There are a few hypotheses as to the molecular mechanism action of mutants that promote the progression of the disease. Thus, it is proposed that ALS pathogenesis may be caused by glutamate excitotoxicity

(Shaw and Ince, 1997), mitochondrial dysfunction (Smith *et al.*, 2017), impaired structure or transport in axons (De Vos *et al.*, 2007) and oxidative stress (Barber and Shaw, 2010).

ALS is a multifactorial disease with both environmental and genetic risk factors. While smoking is the only firmly established environmental factor, there are upwards of 25 genes associated with the disease (Oskarsson *et al.*, 2015; Nguyen *et al.*, 2018). Mutations in *C9orf72* gene account for 30 to 40 % of fALS in USA and Europe. *SOD1* mutations contribute up to 20 % of fALS worldwide. Other genes commonly associated with the risk of ALS are *TARDBP* encoding TAR DNA binding protein 43 (TDP-43), *FUS* encoding fused in sarcoma (FUS) protein binding DNA/RNA which regulates gene expression, *ANG* encoding for a mediator of blood vessel formation or *OPTN* encoding optineurin with a role in Golgi complex maintenance (Zarei *et al.*, 2015). More genes are said to be ALS risk factors, and since there is no full consensus on the cause of the disease, more important factors may still be undiscovered.

1.3.2. Primary lateral sclerosis background

Primary lateral sclerosis (PLS; OMIM %611637) is considered a part of the ALS pathological spectrum. Spiller (1905a; 1905b) was the first to describe PLS. Since then an ongoing debate was had as to the nosology of the disorder and currently PLS is described as a part of the ALS spectrum disease, characterised by spinobulbar spasticity caused by UMN degeneration. PLS is an idiopathic disorder with an adult-onset and accounts for 1-3 % of MNDs. Unlike ALS, PLS does not significantly affect LMNs and has a longer survival time (Finegan *et al.*, 2019). In a study by de Vries *et al.* (2017) 3.3 % of 181 patients with PLS were shown to also have FTD.

Pringle *et al.* (1992) established the clinical features and diagnostic criteria of PLS, based on 8 patients. The clinical features include spasticity, slight weakness in the lower limbs, adult-onset, progressive course, duration of longer than 4 years and pseudobulbar symptoms (Pringle *et al.*, 1992; Kuipers-Upmeijer *et al.*, 2001). The four-year minimum was established by Gordon *et al.* (2006) in order to differentiate between PLS and other MNDs, however Bruyn *et al.* (1995) described cases where patients developed ALS, after initial PLS diagnosis, up to 27 years later. Pathological features include degeneration of pyramidal cells of Betz in the primary motor cortex, which is clearer in PLS than ALS cases. Pyramidal cells of the precentral gyrus are reduced in size in both disorders (Hudson *et al.*, 1993).

The genetic basis of PLS is not well understood. Reports of members of one family having ALS and others PLS exist, but no common genetic risk factor was found (Praline *et al.*, 2010). Mutations in gene *C9orf72* were found in 1 out of 110 PLS patients and 84 out of 1422 ALS patients in a study by van Rheen *et al.* (2012). Mutations in genes *SPG7* and *TBK1* were reported in patients affected by familial PLS (Yang *et al.*, 2016; Gomez-Tortosa *et al.*, 2017). Research on genetic risk factors for PLS needs to be done as it is a serious heterogeneous disease. Many of its features are shared with ALS and therefore further I will discuss ALS, pointing out data specific for PLS.

1.3.3. Epidemiology

The epidemiological data allow, among others, for better distribution of the budget for disease treatment, better understanding of the disease evolution or environmental factors involved in its pathogenesis. The mean age of onset of ALS is between 50 and 65 years and there seems to be an increased risk of ALS in the male to female ratio of approximately 1.5:1 (Zarei *et al.*, 2015). PLS similarly has a median onset age of 50.5 years (Pringle *et al.*, 1992). This information, as well as knowing other possible risk factors such as environmental (smoking, physical activity, exposure to toxic metals, radiation or diet) or genetic, is important in introducing the best strategies to treat the illness with all means available.

1.3.3.1. Prevalence

Arthur *et al.* (2016) estimated the prevalence of ALS at 222,801 cases worldwide in 2015. As mentioned before, men are more at risk of developing ALS than women. Also there exist geographical *loci* where ALS is even 50 to 100 times more prevalent. These include parts of Japan, Guam, Kii Peninsula of Japan, and South West New Guinea. Environmental factors are meant to be the cause of such geographical concentration of ALS cases (Zarei *et al.*, 2015).

Crude prevalence results for Europe range from 1.1/100,000 in Serbia (Alcaz *et al.*, 1996) to 8.2/100,000 individuals in the Faroe Islands (Joensen, 2012). The median prevalence was 5.4/100,000 individuals (Chiò *et al.*, 2013). Estimated prevalence from prospective studies was higher (median of 7.89/100,000 individuals) than from retrospective studies (median of 4.04/100,000 individuals). Prospective studies tend to be less biased than retrospective ones, taking into account outcomes during the study period.

In retrospective studies the outcome is related to the exposures to potential risks looking backwards.

The aging population is supposed to contribute to a 69 % increase of ALS prevalence worldwide by the year 2040 according to Arthur *et al.* (2016). These data from 2015 which estimated the prevalence of ALS to be equal to 222,801, predict 376,674 cases of ALS worldwide by 2040. This increase is predominantly the contribution of developing nations where there is a predicted increase of older individuals from 9 % to 16 % between 2015 and 2040. Yet it may be an underestimate since improving healthcare and treatments may prolong life expectancy of patients and therefore further increase the prevalence.

Prevalence of PLS is much less investigated than that of ALS, however it is estimated at 10 to 20 cases per million (Brugman and Wokke, 2004). Taking under consideration that PLS is much rarer than ALS this estimate appears high, and is likely due to a higher survival time for PLS diagnosed patients.

1.3.3.2. Incidence

Incidence of ALS was estimated at 2.7/100,000 person-years (Chiò *et al.*, 2013). Just as with prevalence, the lowest incidence was seen in Serbia (0.5/100,000 person-years) and the highest in Faroe Islands (3.6/100,000 person-years) in Europe. 2.08 was the median incidence rate (/100,000 person-years) for all of Europe, while for USA the number was 1.8. Incidence rate (/100,000 person-years) for men was higher with 3.0 in comparison to women with 2.4 (Figure 1.10; Logroscino *et al.*, 2009). The disease occurrence after the age 80 years undergoes rapid decrease. This may be due to a reduced number of susceptible individuals after that age, yet it may also be due to diagnostic difficulties. Short life expectancy might hamper the possibility of positive diagnosis.

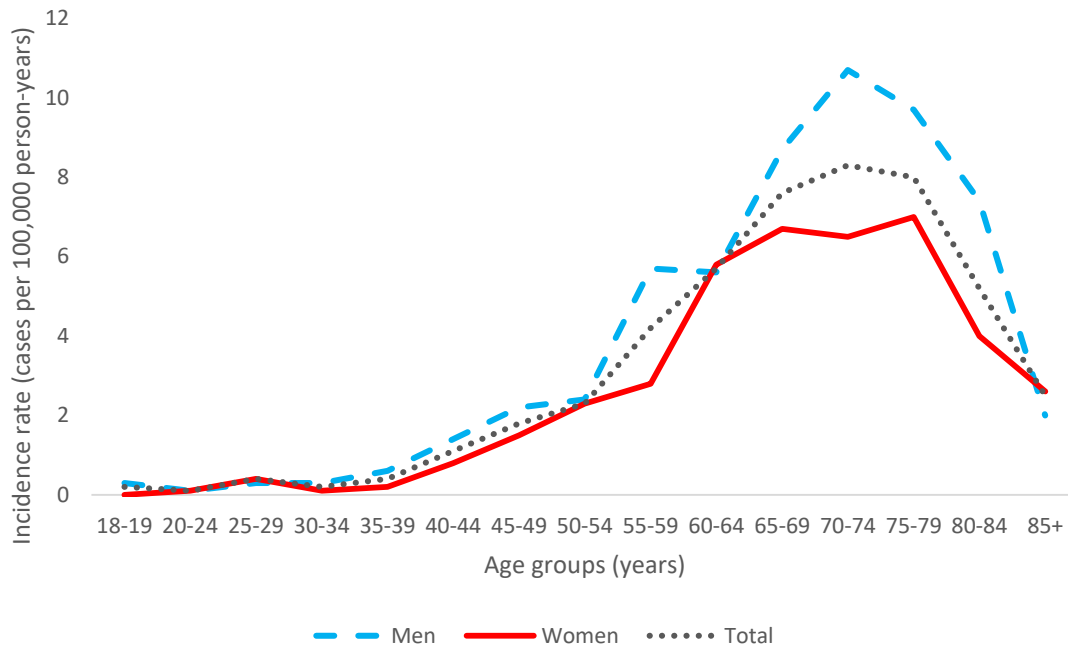


Figure 1.10. Incidence of amyotrophic lateral sclerosis per 100,000 person-years.

Age groups of 18 year olds and older with distinction between men, women and total incidence (data from Logroscino *et al.*, 2009).

Logroscino *et al.* (2009) propose that the limb-onset ALS may contribute to higher incidence of ALS in men. This is due to the observation that limb-onset ALS was twice as common among men. Also limb-onset ALS was much more common than bulbar onset ALS in men, while the difference between the two patterns was much lower in women. They speculate that gender related environmental risks may play a role in this distinction as trauma and strenuous physical activity, proposed as ALS risk factors, are more common in men due to occupational hazards and sports activities.

Arthur *et al.* (2016) predict an increase in ALS incidence around the world of more than 31 % by year 2040. With the greatest increase in Libya (117 %) and Iran (112 %). Incidence in the USA is meant to increase by 34 % and by 20 % in Europe.

The incidence of PLS is very rare and estimated at 1 per 10 million per year (Brugman and Wokke, 2004). Rarity of the disease together with diagnostic difficulties makes this estimate uncertain.

1.3.3.3. Amyotrophic lateral sclerosis epidemiology in Spain

Prevalence of ALS in Spain is similar to the one reported for rest of Europe (5.4/100,000). The same can be said for incidence with 1.4/100,000 person-years (Pradas

et al., 2013; Camacho *et al.*, 2018). Incidence rate reaches a peak in 75 to 79 years old patients. As in other studies, men had a higher disease occurrence at 1.6 in comparison to women at 1.2 (/100,000 person-years). The median onset age was 64.3 years and 38 % of cases had bulbar or generalized onset. Median life expectancy was just above 2.5 years since onset. The data presented were retrieved from a prospective study conducted in Catalonia, Spain (Pradas *et al.*, 2013). As with other studies performed in the Mediterranean area, the prevalence and incidence rates are at the lower range of the European results. With an estimated 3,000 people living with ALS in Spain it is still a rare disease but the socioeconomic burden is significant.

The epidemiological data on PLS in Spain are too scarce to provide an incidence or prevalence estimate.

1.3.4. Clinical features and treatment

ALS is characterized by progressive degeneration of UMNs in the cerebral cortex and LMNs in the medulla and anterior horn of the spinal cord. Muscle weakness that follows, usually leads to a respiratory failure and death within 2-3 years from onset (Coupé and Gordon, 2013). The median survival time is 6 months for 25 % of patients and 12 months for another 25 %. Only about 5 % of patients live 20 years or longer. Predicting the survival time for patients is therefore difficult. Longer survival is associated with limb-onset, young age, stable weight, longer time between onset and diagnosis, higher breathing capacity and motor function (Gordon *et al.*, 2012).

Mean \pm standard deviation (SD) disease duration for PLS was 11.2 ± 6.1 years since onset (Tartaglia *et al.*, 2007). This included patients who died and who lived until the end of the study. Mortality rate was 33 % for PLS patients and 89 % for ALS patients in the same study.

ALS, as the most common MND, encompasses PLS and PMA which make up 1-3 % and 10 % of MNDs respectively (Gordon, 2013). Limb-onset ALS presents both UMN and LMN signs in the limbs. Bulbar onset ALS starts with such symptoms as slurred speech and dysphagia with weakening of limbs later in the disease. PLS shows UMN involvement, while PMA has only LMN involvement (Zarei *et al.*, 2015). Limb-onset ALS is responsible for 70 % of cases among patients and bulbar onset ALS for 25 %.

LMN degeneration causes cramps, wasting of the muscle, fasciculation and weakness while UMN degeneration causes hyperreflexia, twitching, spasticity, modest weakness and emotional lability. The most common limb-onset ALS starts with unilateral

and focal symptoms. Some of the early manifestations are: difficulty walking, loss of dexterity, problems with lifting hands above the head, foot drop, and atrophy of hands, forearms, shoulder, proximal thigh or distal foot muscles. Motor neuron degeneration may eventually lead to a complete loss of limb function with the patients not being able to walk or move their arms, leaving them fully dependent on caregivers. Bulbar onset ALS starts with dysarthria and dysphagia which can later develop to anarthria. Tongue fasciculation and atrophy are characteristic of bulbar onset ALS. Orthopnoea and exertional dyspnea are early symptoms of respiratory difficulty. Approximately 15 % of ALS patients develop FTD and close to 50 % have cognitive impairment. They also suffer from emotional symptoms such as depression or anxiety. Pain is associated with almost 70 % of patients and often caused by bedsores and the inability to move, as well as musculoskeletal conditions such as muscle cramping and spasticity (Coupé and Gordon, 2013; Zarei *et al.*, 2015).

A study comparing ALS and PLS (Tartaglia *et al.*, 2007) showed a statistically significant difference in the age of onset with 59.1 ± 13.1 and 54.6 ± 10.9 years, respectively (mean age \pm SD; $p = 0.009$). 47 % of PLS patients exhibited stiffness *versus* 4 % of ALS patients, at first presentation at the clinic. Dysphagia was a symptom among 8 % of ALS patients and none of the PLS patients. Other symptoms more common in ALS patients were weakness, cramps, fasciculations and limb wasting. At follow-up a significant difference was seen in limb wasting with only 2 % of PLS patients and 100 % of ALS patients affected. Bulbar symptoms and dementia were more common symptoms among ALS patients and cortical signs among PLS patients.

Diagnosis of ALS is done using the El Escorial criteria established by the World Federation of Neurology in 1994 and revised in 2015 (Ludolph *et al.*, 2015). Electromyogram is a very helpful tool in ALS diagnosis confirming widespread LMN degeneration (Zarei *et al.*, 2015).

Thus far no cure for ALS exists. Riluzole approved by the U.S. FDA in 1995 and by the European Medicines Agency (EMA) in 1996 was the first systemic therapy drug for ALS. Riluzole is a non-competitive NMDA receptor antagonist and inactivates voltage-gated sodium channels. It is supposed to reduce glutamate excitotoxicity. Edaravone, approved by the FDA in 2017, is an antioxidant, reducing reactive oxygen species (ROS) and thus oxidative stress. Other compounds have been tested for ALS treatment and they fall into the following pathophysiological categories: antiapoptotic, anti-inflammatory, ant glutamatergic, antioxidant or anti-aggregation (Zarei *et al.*, 2015).

PLS patients are not included in riluzole nor edaravone trials as according to the El Escorial criteria only “definite” or “probable” ALS patients can partake (Finegan *et al.*, 2019). This creates a difficult situation for people affected by this disease and revision of those requirements might be necessary as PLS and ALS may have common pathogenesis.

1.3.5. Amyotrophic lateral sclerosis pathology

Pathogenesis in ALS seems to be multifactorial with many mechanisms, rather than a single one, contributing to neurodegeneration (Figure 1.11). The mechanisms related to ALS pathogenesis are RNA processing, mitochondrial dysfunction, glutamate excitotoxicity, oxidative stress, protein aggregation, neurofilaments (NFs) accumulation and neuroinflammation. The exact way by which ALS neurodegeneration occurs is not clear, however the above-mentioned mechanisms may interact with each other causing the disease.

Pathogenesis in PLS is much less understood. However, the mechanisms involved in neurodegeneration in PLS may be similar to those in ALS.

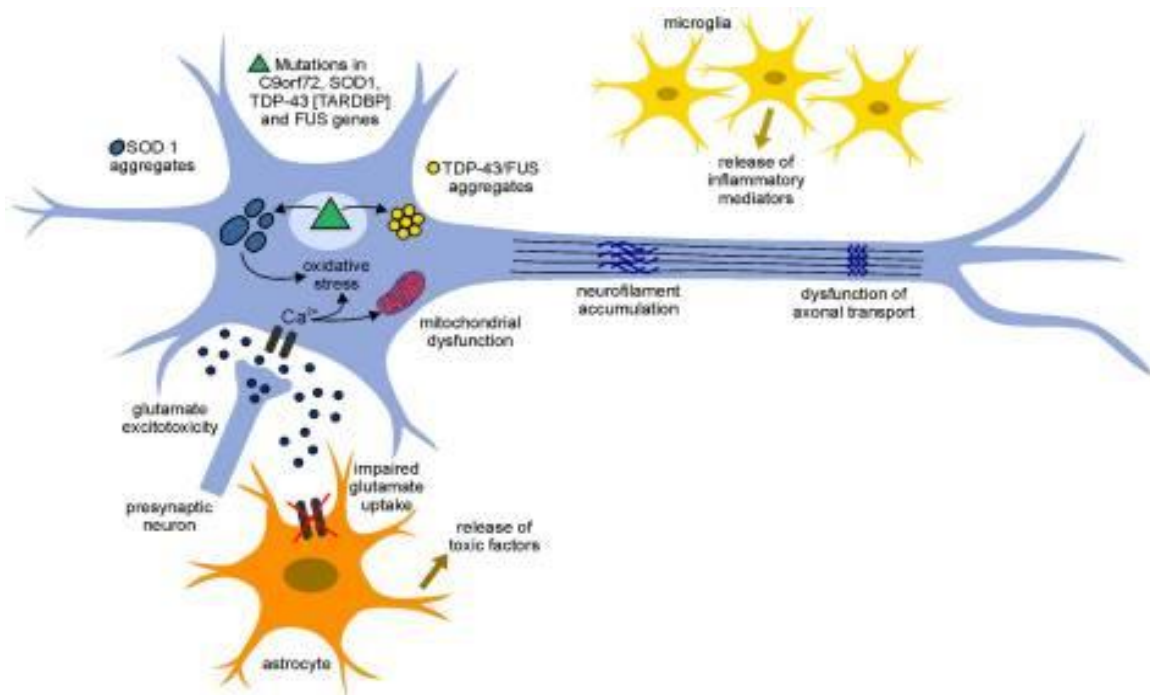


Figure 1.11. Amyotrophic lateral sclerosis pathogenesis mechanisms.

Figure 1.11. Amyotrophic lateral sclerosis pathogenesis mechanisms (continued).

Glutamate excitotoxicity caused by impairment of astrocyte uptake of neurotransmitter causes Ca^{2+} influx into motor neurons. Mitochondrial dysfunction impedes calcium removal and thus oxidative stress may be caused through calcium-dependent enzymatic pathways. Misfolding of mutant SOD1, C9orf72, TDP-43 and FUS causes them to aggregate which exacerbates oxidative stress, mitochondrial dysfunction, NFs accumulation and thus axonal transport. Glial cells contribute to motor neuron death through release of pro-inflammatory cytokines and neurotoxins. Figure from Bonafede and Mariotti, 2017 licensed under CC BY 3.0.

1.3.5.1. RNA processing

Among genes associated with ALS, some code for RNA-binding proteins (*FUS*, *TARDBP*). Pathology of TDP-43 encoded by *TARDBP* is common in ALS patients. Both *FUS* and TDP-43 are involved in pre-mRNA splicing, RNA transport and RNA translation. Among other genes associated with ALS and involved in RNA processing pathway are *ANG*, *ELP3*, *SETX*, *SMN* and *C9orf72*. They are involved in RNA transcription regulation, transcription, splicing, post-transcriptional processing/editing, transport, translation regulation and degradation. When *C9orf72* has an inappropriate number of hexanucleotide GGGGCC repeats, its pre-mRNA may attach to RNA-binding proteins in the nucleus, thus, making them unavailable for processing of other mRNAs (van Blitterswijk and Landers, 2010; Morgan and Orrell, 2016).

Substrates of TBK1, whose mutants are associated with PLS, have many substrates of their own, among which also proteins involved in RNA processing (Gomez-Tortosa *et al.*, 2017).

1.3.5.2. Mitochondrial dysfunction

Damage and dysfunction of mitochondria are common in many neurodegenerative diseases. Mitochondria are important in energy metabolism, apoptosis, lipid biosynthesis and calcium homeostasis. They also produce ROS at high levels which may contribute to oxidative stress. SOD1 is present in mitochondria at lower levels than SOD2, yet misfolded SOD1 may affect the function of these organelles. Another gene associated with ALS and encoding for a mitochondrial protein is *CHCHD10*. It is located in the intermembrane space and cristae junctions. Differential motor neuron vulnerability, hyperexcitability or

fasciculations are among ALS characteristics that may be explained by mitochondrial dysfunction.

ALS pathogenesis has been hypothesized to be directly linked with disruption of mitochondrial bioenergetics, structure and dynamics. Calcium buffering is also an important feature disrupted in ALS. The presence of mutant SOD1 in ALS patients has been shown to be associated with loss of Ca^{2+} binding proteins in motor neurons. This reduces Ca^{2+} uptake and increases sensitivity to excitotoxicity (Morgan and Orrell, 2016; Bonafede and Mariotti, 2017; Smith *et al.*, 2017).

In adult-onset PLS the mitochondrial function was affected under reduced glucose conditions by mutant *SPG7* associated with the disorder (Yang *et al.*, 2016).

1.3.5.3. Glutamate excitotoxicity

Glutamate is one of the main neurotransmitters in the central nervous system (CNS). Synthesized at the presynaptic neuron it is an excitatory activator of postsynaptic receptors. It acts on NMDA and α -amino-3-hydroxyl-5-methyl-4-isoxazole-propionate (AMPA) receptors activating them, which in turn leads to postsynaptic membrane depolarization and activation of voltage-dependent calcium channels and influx of Ca^{2+} . Glutamate removal from the synaptic cleft is done by excitatory amino acid transporters (EAATs), glial and neuronal proteins. If the activation of postsynaptic receptors is too long or excessive, it may lead to a state of excitotoxicity and death of neurons. Superfluous intracellular calcium may have detrimental effects on mitochondria and formation of ROS, which may lead to mitochondrial dysfunction and oxidative stress and eventually cell death.

Glutamate excitotoxicity was one of the first processes implicated in ALS. Patients with this disease were shown to have a reduced level of EAAT2, one of the most important cell transporter protein involved in keeping the glutamate level below excitotoxic. Excessive influx of Ca^{2+} , excitation of motor neurons and initiation of destructive biochemical processes in them, are all pathophysiological mechanisms in both fALS and sporadic ALS (sALS) (Morgan and Orrell, 2016; Bonafede and Mariotti, 2017).

1.3.5.4. Oxidative stress

Aerobic metabolism produces ROS, which are effectively removed in natural processes in the cell. When the cell is not capable of removing these free radicals to ensure homeostasis, ROS accumulate and cause damage to proteins, lipids, DNA and RNA. SOD1 is an enzyme whose function is to break down superoxide radicals after binding copper and

zinc molecules. Thus mutations in *SOD1*, associated with ALS, can reduce its activity promoting cytotoxic damage caused by charged oxygen molecules. It has been proposed that mutations in *SOD1* may, not only cause a loss-of-function of the enzyme, but to invert its activity producing superoxide radicals and causing oxidative stress. ALS patients were shown to have higher levels of ROS and damage due to oxidative stress in CSF, urine and serum (Barber and Shaw, 2010; Bonafede and Mariotti, 2017).

1.3.5.5. Protein aggregation

Aggregated mutant SOD1 was found in tissues of patients with sALS and fALS. Protein aggregates are a very common characteristic of many neurodegenerative diseases. Such inclusions are typically formed by oligomerization of misfolded protein and are supposed to cause toxicity leading to cell death. TDP-43 and FUS were also found to form aggregates in patients with ALS. Mutations in 3'-untranslated region (UTR) of *TARDBP* seem to cause an aberrant localization and overexpression of TDP-43, which accumulates in the cytoplasm, while being a nuclear protein. Aggregates of TDP-43 have been seen in approximately 80 % of the disease cases. Other components have also been observed in protein inclusions in tissues of ALS patients. Chaperones, NFs, ubiquitin, ubiquitinated proteins and mitochondrial proteins are some of them. This aggregation of proteins may further damage proteasome machinery, increasing accumulation of proteins by impairing their degradation. This leads to motor neuron death and degeneration (Bonafede and Mariotti, 2017).

Patients affected by PLS with mutation in *TBKI* have not been studied with regards to the mechanisms of pathogenesis (Gomez-Tortosa *et al.*, 2017) but are possibly affected by TDP-43 neuropathology as previously described for *TBKI* loss-of-function mutation (Van Mossevelde *et al.*, 2015).

1.3.5.6. Neurofilaments accumulation

NFs are intermediate filaments particularly abundant in axons. During development they are important for radial growth of axons. They also play a role in axon calibre maintenance and transmission of electrical signals in axons. Accumulation of NFs within cell bodies and proximal axons is one of the hallmarks of ALS. In the CNS, NFs are made up of heteropolymers of heavy, medium and light subunits (NF-H, NF-M, NF-L respectively) and α -internexin. In the peripheral nervous system (PNS) α -internexin is replaced by peripherin (Yuan *et al.*, 2012). Phosphorylation of NF-H and NF-M in the axon

is inversely correlated to NF transport. Abnormal phosphorylation of these filaments may be due to mutations in genes encoding NFs which are associated with ALS. These mutations may therefore affect the axonal transport and aggregation of NFs, also impeding transport of mitochondria along the axon (Bonafede and Mariotti, 2017). It has been shown that overexpression of NF-L causes accumulation of NFs in the cell body and proximal axon and degeneration of motor neurons in the ventral horn of the spinal cord (Xu *et al.*, 1993). At the same time overexpression of NF-L and NF-H seems to slow down the progression of the disease. Both these studies were done using a *SOD1*^{G93A} transgenic mouse model of ALS. It is still unclear how exactly NF accumulation affects ALS (Bonafede and Mariotti, 2017).

1.3.5.7. Neuroinflammation

Activated microglia, which are the resident macrophages of the CNS, or astrogliosis, characterize the neuroinflammatory response in many neurodegenerative diseases. Glial cells are especially important in ALS, as disruption of their communication with motor neurons may cause neuronal death. It has been reported that mutant *SOD1* motor neurons surrounded by wild-type glial cells do not develop a pathological phenotype, while wild-type motor neurons surrounded by mutant *SOD1* expressing glia do show signs of pathology (Clement *et al.*, 2003).

Microglia have an anti-inflammatory and neuroprotective phenotype at the initial stages of ALS. Later on, when motor neurons and astrocytes release misfolded proteins, microglia are activated and they become proinflammatory and neurotoxic. Microglia can be activated to M1 or alternatively activated to M2. The M1 cause motor neuron death through proinflammatory cytokines, ROS, and neurotoxins, while M2 produce anti-inflammatory cytokines and neurotrophic factors that protect motor neurons. This is especially important as microglia expressing mutant *SOD1* become M1 activated. Microglia show M2 phenotype in the initial stage of ALS and M1 phenotype as the disease progresses (Bonafede and Mariotti, 2017).

The most abundant cells of the CNS, astrocytes, are very important in supporting neurons. They maintain homeostasis, recycle neurotransmitters and provide nutrients for other cells. Reactive astrogliosis is a term used for morphological and functional changes in astrocytes in response to neurological diseases and has been associated with neurodegenerative diseases and ALS progression (Pekny and Pekna, 2014). Astrocytes remove glutamate from the synaptic cleft via EAAT2 glutamate receptors, preventing

excitotoxicity. Patients with ALS have been observed to have a downregulated EAAT2 in astrocytes (Howland *et al.*, 2002). Zhao *et al.* (2013) have studied astrocytes from post-mortem tissues of ALS patients showing that 22 genes expressing proinflammatory cytokines, complement pathway proteins and chemokines were upregulated. These could be involved in degeneration of affected neurons.

1.3.6. Etiology

1.3.6.1. Environmental factors

Some of the environmental risk factors for ALS include smoking, military service, exposure to lead, pesticides, physical activity, head trauma, electromagnetic radiation, low body mass index, statin treatment or consumption of β -N-methylamino-l-alanine (BMAA) (Oskarsson *et al.*, 2015). BMAA, produced by cyanobacteria in the roots of cycad *Cycas micronesica*, is ingested by people in the areas where they consume cycad seeds. Inability to prevent BMAA accumulation is supposed to increase the risk of ALS.

Men are 1.5-fold more likely to develop ALS which is reflected in the epidemiological data. This may be due to many different factors, but a genetic one seems less likely than an environmental. Smoking is an environmental factor which is connected to ALS, especially in smokers who started at a young age. It is not clear how smoking affects the disease but oxidative stress, toxic substances and inflammation may be the culprits. Also athletes have been observed to have a higher risk of developing ALS, suggesting that physical activity or injury might have a role to play. Exposure to agricultural chemicals such as pesticides or fertilizers and to heavy metals like lead has also been associated with ALS (Oskarsson *et al.*, 2015; Zarei *et al.*, 2015).

Men are more likely to develop PLS, however no environmental factors are known. This is due to lack of accumulated knowledge on the topic and rarity of the disorder.

1.3.6.2. Genetic basis of amyotrophic lateral sclerosis and primary lateral sclerosis

Classical and high-throughput DNA sequencing techniques allowed for identification of many genes associated and causative of ALS. Genetic risk factors have been identified for both fALS and sALS. The most common genes associated with the disease are: *C9orf72*, *SOD1*, *TARDBP* and *FUS*. They contribute to approximately 69 % of all fALS cases and 9 % of sALS cases (Figure 1.12; Morgan and Orell, 2016).

Genes associated with PLS are *SPG7*, *TBK1* and *PSEN1* (Yang *et al.*, 2016; Gomez-Tortosa *et al.*, 2017; Jazi *et al.*, 2019). A locus for PLS was also identified at Chromosome 4ptel-4p16.1 (Valdmanis *et al.*, 2008). Mutations in two more genes (*ALS2* and *ERLIN2*) cause juvenile PLS. Brugman *et al.* (2007) suggest that adult-onset PLS is not associated with *ALS2* gene mutations.

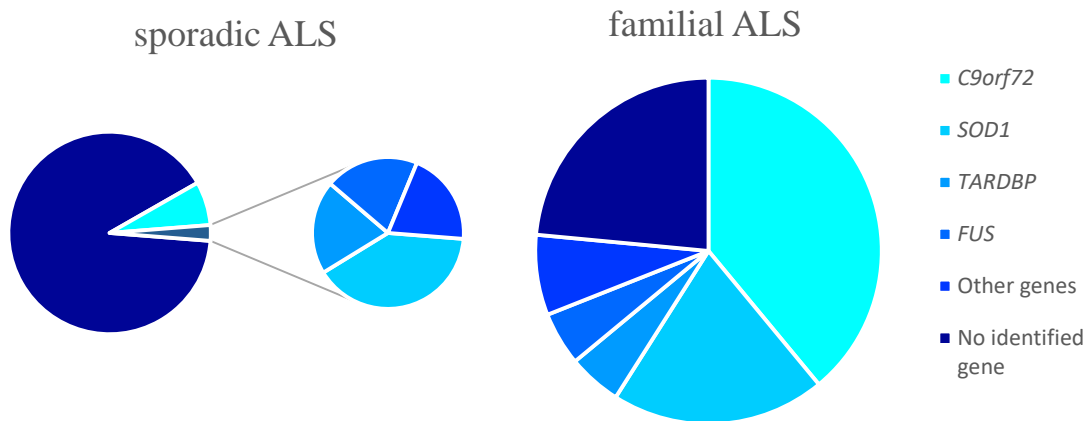


Figure 1.12. Approximate contribution of mutations in sALS and fALS. The sALS representing approximately 90 % and the fALS 10 % of all ALS patients (data from Morgan and Orell, 2016).

C9orf72 gene

Although the function of the protein encoded by *C9orf72* is not known, it is thought to be involved in RNA processing. It is the gene most frequently linked to both fALS and sALS, as well as FTD in Europe. Approximately 40 % of fALS cases can be associated with this gene, however it accounts for less than 10 % of fALS in Asia. The pathological trait has an age-dependent penetrance. Before the age of 35 years it is estimated to be 0 %, increasing to 50 % at 60 years and close to a 100 % at 80 years. *C9orf72* has 12 exons and 3 transcripts. A varying number of the hexanucleotide repeat GGGGCC in the promoter region of one of the transcripts, which corresponds to the first intron of the other two, is linked to pathology in ALS (Corcia *et al.*, 2017). The number of repeats is polymorphic and it is not clear what size should be considered pathogenic and which normal. The repeat size in unaffected individuals seems to vary between 2 and 24. ALS patients have been reported to possess hundreds, even thousands, of such repeats. Although smaller sizes have

also been reported, with 50 being the smallest repeat number co-segregating with the disease (Gijselinck *et al.*, 2016).

In a review, Nguyen *et al.* (2018) point to three main pathological mechanisms of *C9orf72*. The first one being loss-of-function and haploinsufficiency, where the repeat expansion would suppress gene expression. The second is RNA toxicity. This hypothesis suggests the RNA transcripts of the expansion would sequester RNA-binding proteins and hinder RNA processing. The third is proteotoxicity from aggregates of dipeptide repeat (DPR). Toxic poly-GA, poly-GP, poly-GR, poly-PR and poly-PA peptides are formed by repeat-associated non-AUG (RAN) translation of the hexanucleotide repeat GGGGCC or GGCCCC RNA. Toxic DPR inclusions have been recorded in the hippocampus, frontotemporal cortex and cerebellum, causing neurodegeneration.

C9orf72 hexanucleotide repeats were found only in approximately 1 % of PLS patients *versus* 6 % of ALS patients in a study by van Rheenen *et al.* (2012), which does not exclude the involvement of the RNA processing pathway as a possible pathogenesis mechanism in PLS.

SOD1 gene

SOD1 encodes a homodimeric metalloenzyme involved in changing O_2^- into O_2 and H_2O_2 , in the Fenton reaction. To date over 150 mutations in this gene have been described and it is the second most common ALS causative gene after *C9orf72*. Approximately 20 % of fALS and 1-4 % of sALS are caused by mutations in *SOD1*. Most of these mutations are of missense type and follow an autosomal dominant form of inheritance.

The mechanism by which *SOD1* mutations cause ALS is not fully understood. Different mutations may show variation in phenotype caused, age of onset, severity, progression and duration of the disease. Autosomal dominant mutation *SOD1*^{D90A} was observed to be recessive in a Scandinavian population. A shorter survival time has been reported for patients carrying the *SOD1*^{A4V} mutation. It was also reported to have a 91 % penetrance (Pasinelli and Brown, 2006).

Oxidative phosphorylation by-products, superoxide radicals which may transform into more toxic ROS are broken down by *SOD1*. Mutations in *SOD1* may contribute to forming protein aggregates by misfolding, oxidative stress, mitochondrial dysfunction, inflammation, glutamate excitotoxicity and axonal transport abnormalities (Chen *et al.*, 2013). It has been proposed that *SOD1* mutations may be involved in ALS pathology

through gain rather than loss-of-function (Morgan and Orrell, 2016). To date the most commonly used mouse model of ALS is SOD1^{G93A} mutant.

TARDBP gene

TDP-43 encoded by *TARDBP* gene is a ribonucleoprotein involved in DNA/RNA machinery. Sreedharan *et al.* (2008) reported an involvement of *TARDBP* mutations in fALS. Since then *TARDBP* has become one of the most important ALS genes accounting for approximately 5 % of fALS and < 1 % of sALS cases (Corcia *et al.*, 2017). A great majority of mutations in this gene are located in the glycine-rich C-terminal region in the sixth exon. This region plays a role in interactions between proteins. ALS patients with mutations in *TARDBP* seem to present an autosomal dominant inheritance with predominant limb-onset in Caucasian and bulbar onset in Asian ALS populations. Ubiquitinated TDP-43 is part of the cytoplasmic inclusions not only in ALS but also FTD, Huntington disease (HD), Parkinson's disease (PD) and AD (Da Cruz and Cleveland, 2011). TDP-43 ubiquitin-positive inclusions are not reported in ALS patients with *SOD1* mutations. It is not clear whether it is due to gain or loss-of-function that TDP-43 is related to ALS pathogenesis. Since the protein is involved in transcription, splicing, mRNA transport and other nucleus processes both gain of toxic and loss-of-function can cause significant alterations in RNA metabolism (Chen *et al.*, 2013).

FUS gene

Searching for a *locus* with a high logarithm of odds (LOD) within genes with similar domains as in *TARDBP*, the *FUS* gene was identified as a risk factor for ALS (Vance *et al.*, 2009). Autosomal dominant and recessive inheritance was reported for *FUS* mutations. Just as with *TARDBP*, *FUS* also accounts for approximately 5 % of fALS and < 1 % of sALS cases and is involved in DNA/RNA machinery. Mutations in *FUS* related to ALS seem to be located exclusively in exon 13. They cause aberrant distribution of this nuclear protein in cytoplasm (Corcia *et al.*, 2017). Age of onset is typically young although it can vary between 26 and 80 years with the youngest patient recorded at 11 years. *FUS* related ALS is predominantly affecting LMNs and does not involve the bulbar region nor cognitive impairment (Chen *et al.*, 2013). *FUS* is involved in gene expression and mRNA splicing among other nuclear processes, thus its mechanisms of pathogenesis are similar to those of *TARDBP* (Morgan and Orrell, 2016).

Primary lateral sclerosis genes

Compound heterozygous mutations in *SPG7* were found to be associated with PLS in a family affected by the disorder. The mutations were *SPG7*^{L695P} and *SPG7*^{L743T}, which cosegregated with the disease in the family in an autosomal recessive fashion (Yang *et al.*, 2016). The affected members had both of these mutations while the unaffected parents and siblings carried only one. Using the SeaHorse metabolic analyzer Yang *et al.* (2016) showed an alteration in mitochondrial function under reduced glucose conditions affected by *SPG7* mutations.

A Spanish family with PLS/dementia phenotype was identified as having *TBK1* gene mutation (*TBK1*^{R573G}) segregating with the disease (Gomez-Tortosa *et al.*, 2017). No neuropathological studies were performed to assess the impact of the mutation. Yet it is located at the dimerization domain which may cause protein instability or prevent *TBK1* substrate binding. Also TDP-43 neuropathology has been previously associated with a mutation in *TBK1* (Van Mossevelde *et al.*, 2015).

Jazi *et al.* (2019) have identified mutations *PSEN1*^{A431E} and *PSEN1*^{L381V} as associated with PLS. The former identified in two Mexican-American brothers and the latter in a Vietnamese man. The two brothers developed spastic quadriparesis after previous FTD diagnosis and the Vietnamese man developed progressive spastic quadriparesis after cognitive problems since childhood. All developed bulbar dysfunction symptoms. The early upper extremity involvement and bulbar dysfunction allowed for possible exclusion of hereditary spastic paraparesis (HSP) as the disease affecting the patients.

Other genes

Rarer genes related to ALS include *OPTN*, *ANG*, vehicle-associated membrane protein B (*VAPB*), valosin-containing protein (*VCP*), alsin (*ALS2*), spatacsin (*SPG11*), heavy chain neurofilament (*NEFH*) or X chromosome linked ubiquilin-2 (*UBQLN2*) (Morgan and Orrell, 2016). There is a great number of genes reported as causative or risk factors of fALS and many of them are rare cases (Table 1.3). This shows the heterogeneous pathology of ALS and the possibility of discovering new genes involved in the disease.

Table 1.3. Genes associated with fALS and their inheritance and clinical features.

fALS	<i>Locus</i>	<i>Gene</i>	Inheritance	Clinical features
ALS1	21q22.1	<i>SOD1</i>	AD, AR	fALS
ALS2	2q33	Unknown	AR	Juvenile ALS
ALS3	18q21	Unknown	AD	fALS
ALS4	9q34	<i>SETX</i>	AD	Juvenile ALS
ALS5	15q15-21.1	<i>SPG11</i>	AR	Juvenile ALS
ALS6	16q12	<i>FUS</i>	AD–AR	fALS
ALS7	20p13	Unknown	AD	fALS
ALS8	20q13.33	<i>VAPB</i>	AD	fALS
ALS9	14q11	<i>ANG</i>	AD	fALS
ALS10	1p36.22	<i>TARDBP</i>	AD	fALS
ALS11	6q21	<i>FIG4</i>	AD	fALS
ALS12	10p15	<i>OPTN</i>	AD, AR	fALS
ALS13		<i>ATXN2</i>	AD	fALS
ALS 14	9p13	<i>VCP</i>	AD	fALS
ALS15	Xp11	<i>UBQLN2</i>	X-linked	fALS
ALS16	9p13.3	<i>SIGMAR1</i>	AR	fALS
ALS 17	3p11	<i>CHMP2B</i>	AD	fALS
ALS18	17p13	<i>PFN1</i>	AD	fALS
ALS 19	2q34	<i>ERBB4</i>	AD	fALS
ALS 20	12q13	<i>HNRNPA1</i>	AD	fALS
ALS 21	5q31	<i>MATR3</i>	AD	fALS
ALS 22	2q35	<i>TUBA4A</i>	AD	fALS
ALS	12q24	<i>ADO</i>	AD	fALS
ALS	9q34	<i>GLE1</i>	AR	fALS
ALS	20q13	<i>SSI8L1</i>		fALS
FTD–ALS1	9p21	<i>C9ORF72</i>	AD	ALS, FTLD, ALS–FTD
FTD–ALS2	22q11	<i>CHCHD10</i>	AD	ALS, FTLD, ALS–FTD
FTD–ALS3	5q35.3	<i>SQSMT1</i>	AD	ALS, FTLD, ALS–FTD
FTD–ALS4	12q14.2	<i>TBK1</i>	AD	ALS, FTLD, ALS–FTD

(data from Corcia *et al.*, 2017).

1.3.7. ADP-ribosylhydrolases and ADP-ribosyltransferases

ADP-ribosylhydrolases (ARHs) and ADP-ribosyltransferases (ARTs) function opposite each other in ADP-ribosylation cycle although some ADP-ribosylation events are considered irreversible under the current state of knowledge (Rack *et al.*, 2020). ADP-ribosylation is a posttranslational modification (PTM) of proteins with great importance for cell regulation. ARTs are involved in transfer of single or multiple ADP-ribose units from β -nicotinamide adenine dinucleotide (β -NAD⁺) to a substrate. The transfer of ADP-ribose onto a protein can be divided into mono-(ADP-ribosyl)ation (MARylation) and poly-(ADP-ribosyl)ation (PARylation). Different ARTs may be responsible for MARylation and PARylation and different hydrolases cleave mono-ADP-ribose and poly-ADP-ribose. Some members of ARHs have been described as associated with neurodegeneration (Sharifi *et al.*, 2013; Danhauser *et al.*, 2018; Mashimo *et al.*, 2019). McGurk *et al.* (2018) identified a possible misregulation of PARylation in ALS spinal cord motor neurons associated with nuclear PAR polymerases 1 and 2 (PARP1 and PARP2).

The ADP-ribosylation is an intricate PTM induced by a varied array of ARTs and with different chemical linkages (Table 1.4). The *O*-glycosidic linkages, *N*-glycosidic linkages and *S*-glycosidic linkage are known to be a part of this PTM. There are more than 20 different transferases and at least 8 hydrolases described for ADP-ribosylation PTM.

The ADP-ribosylhydrolase family of hydrolases includes three members described in humans (ADPRH, ADPRHL1 and ADPRS (previously ADPRHL2)) and one first found in *Rhodospirillum rubrum* (DraG), a diazotrophic, photosynthetic α -proteobacterium (Ludden and Burris, 1976; Moss *et al.*, 1985). Summarising the current knowledge on ADPRH, ADPRHL1 and ADPRS is important due to their significance for this thesis. The amino acid similarity between ADPRH and ADPRHL1 is 68 % and between ADPRH and ADPRS, 41 %. ADPRS hydrolyses *O*-glycosidic bond in PARylated substrates (Mashimo *et al.*, 2014). ADPRHL1 does not seem to have an enzymatic activity and it is mostly restricted to expression in the heart of vertebrates (Smith *et al.*, 2016). Mutations in ADPRS are known to be associated with neurodegeneration (Danhauser *et al.*, 2018; Mashimo *et al.*, 2019), while ADPRH is involved in tumorigenesis and bacterial response to toxins (Mashimo *et al.*, 2014).

Table 1.4. Summary of ADP-ribosylation intricate array of chemical linkages, transferases, hydrolases and their substrates.

	Modification targets	Examples of known substrates	Transferases	Hydrolases
<i>O</i> -glycosidic linkages	Glutamic/aspartic acid	β -TrCP, GSK3b, LXR α/β , NXF1, PARP1, PARP2, PARP3, PARP5a, PARP5b, PARP10, PARP11, PARP13, PARP16, PCNA	PARP1, PARP2, PARP3, PARP5a, PARP5b, PARP7, PARP10, PARP11, PARP14, PARP16	MacroD1, MacroD2, TARG1
	Aspartic acid	GcvH-L, PARP12	SirTM, PARP6, PARP12	MacroD1, MacroD2, TARG1
	C terminus	Ubiquitin	PARP9	Unknown
	Acylated lysine	OAADPr	Sirtuins	MacroD1, MacroD2, TARG1, ADPRS
	Serine and tyrosine	PARP1, histone H1, H2B, H3, HPF1	PARP1/2:HPF1 complex	ADPRS
	ADP-ribose	2'-OH and 2''-OH	PARP1, PARP2, PARP5a, PARP5b	PARG, ADPRS
	3'/5'-phosphoRNA, 3'/5'-phosphoDNA, 2'-phosphoRNA	tRNA	KptA/TRPT1, PARP1, PARP3, PARP10, PARP11, PARP15	MacroD1, MacroD2, TARG1, ADPRS, PARG, NUDT16

Table 1.4. Summary of ADP-ribosylation intricate array of chemical linkages, transferases, hydrolases and their substrates (continued).

N-glycosidic linkages	Arginine	integrin $\alpha 7$, hemopexin, GRP78/BiP, GS α	hARTC1, hARTC5, cholera toxin	ADPRH
	Lysine	PARP16	PARP16	Unknown
	Diphtamide	EF2	exotoxin A	Irreversible
	Guanine	dsDNA	Pierisin, CARP-1, ScARP	Irreversible
S-glycosidic linkage	Cysteine	PARP8, PARP16	PARP8, PARP16	Unknown

(data from Rack *et al.*, 2020)

ADPRH is a ubiquitous protein found in cytoplasm of both mice and humans (Moss *et al.*, 1992). ADPRH hydrolyses the *N*-glycosidic bond between arginine and ADP-ribose cleaving ADP-ribose from substrate (Oka *et al.*, 2006; Ono *et al.*, 2006). The reaction occurs to be specific to mono-ADP-ribosylated substrates, due to the protein's structure (Rack *et al.*, 2018). Kato *et al.* (2011) observed a shorter G1 phase and increased cell growth in mouse embryonic fibroblasts with depleted ADPRH. Therefore, ADPRH has been implicated in cell proliferation, tumorigenesis, intracellular signal transduction and cell cycle regulation. Mutations in ADPRH such as p.D56N are implicated in altering the active site of the protein and thus its catalytic activity, consequently being implicated in cancer progression (Kato *et al.*, 2015; Rack *et al.*, 2018).

Interestingly ADPRH has been reported to hydrolyse free arginine-ADP-ribose, which can be ADP-ribosylated by cholera toxin secreted by *Vibrio cholerae*. In turn, this free arginine-ADP-ribose seems to be an inhibitor of ADPRS (Drown *et al.*, 2018; Rack *et al.*, 2018). As mentioned previously, ADPRS is involved in neurodegenerative diseases and cleaves *O*-glycosidic serine bonds of ADP-ribosylated substrates. Just as ADPRH it is ubiquitous, and present in cytosol, mitochondria and nucleus (Rack *et al.*, 2020). ADPRS is also implicated in regulation of parthanosis, a form of programmed cell death (Mashimo *et al.*, 2013; Dawson *et al.*, 2017). Also the loss-of-function mutations in ADPRS which are associated with a neurodegenerative disorder have suggested its possible role in

neuroprotection from endogenous ROS (Danhauser *et al.*, 2018; Ghosh *et al.*, 2018). The presence of ADPRS in the mitochondria has not been studied very well, but Niere *et al.* (2012) suggest it is responsible for degradation of mitochondrial matrix-associated PAR. Also Dölle *et al.* (2013) indicate that degradation of O-Acetyl-ADP-ribose (OAADPr) by ADPRS may be important in NAD recycling in mitochondria.

ADPRHL1, which has more similarity with ADPRH than ADPRS, is reported to lack catalytic activity of ADP-ribose cleavage, being restricted to the heart and affecting heart chamber outgrowth and muscle actin filament assembly (Smith *et al.*, 2016). Although, thus far, no protein expression of ADPRHL1 has been reported outside of the heart, RNA expression was found in several other tissues. Amongst them the brain, and especially the cerebral cortex (Uhlén *et al.*, 2015). The catalytic activity of ADPRHL1 is reported to be lost (Oka *et al.*, 2006) due to the lack of key amino acids of the active site, otherwise present in ADPRH (D56, Y263, S264, S269, S270, D302, D304; Smith *et al.*, 2016). Therefore, ADPRHL1 has been described as a pseudoenzyme, a protein which has lost its essential catalytic activity.

Smith *et al.* (2020) suggest actin may be a shared target for ADPRH and ADPRHL1 in *Xenopus* species. ADPRH may be important in cleaving ADP-ribose from actin at residue R177. Nine toxins are known with the transferase activity of ADP-ribosylation of actin monomers at that residue. This in turn provokes a collapse of actin filaments. ADPRH and ADPRHL1 may work in unison to allow for correct actin filament assembly.

1.4. Paroxysmal kinesigenic dyskinesia

1.4.1. Background

Paroxysmal kinesigenic dyskinesia (PKD; OMIM #128200) is the most common paroxysmal movement disorder characterized by a range of involuntary movements triggered by sudden motion. PKD is considered a rare disorder, which disproportionately affects men (Chen *et al.*, 2011).

This rare autosomal dominant neurologic condition was first described as early as 1892 by Shuzo Kure in a Japanese medical journal (Kure, 1892). The typical onset of the disorder is during childhood or adolescence. Recurrent attacks involving chorea, athetosis, dystonic postures or less commonly ballismus are symptoms of the disorder which is often misdiagnosed clinically as epilepsy. PKD is also clinically similar to two allelic disorders: benign familial infantile epilepsy (BFIE) and infantile convulsions and paroxysmal

choreoathetosis (ICCA). The attacks may occur as seldom as once per month to as often as a hundred times per day, lasting up to 5 minutes each. Their severity typically decreases with age (Chen *et al.*, 2011). No loss of consciousness has been reported during the attacks. PKD primary dysfunction is most probably caused by a global dysfunction of the motor neuron network (Méneret *et al.*, 2013).

Thus far the only known cause of this condition are mutations in the proline-rich transmembrane protein 2 (*PRRT2*) gene, which has an autosomal dominant pattern of inheritance. Although the exact function of *PRRT2* has not been identified, it is known to interact with synaptosomal nerve-associated protein 25 (SNAP-25), synaptotagmin-1 (SYT1) and synaptotagmin-2 (SYT2) in the presynaptic membrane of neurons, which are involved in signalling in nerve cells (Valente *et al.*, 2016). Mutations in *PRRT2* may decrease the amount of *PRRT2* available at the membrane and therefore alter the activity of neurons, which in turn may cause the symptoms observed in PKD patients.

At least 17 known mutations in *PRRT2* are associated with forms of paroxysmal dyskinesia. A hotspot for mutations is at c.649 on exon 2 of this 4 exon gene. Three truncating mutations at c.649 account for 85 % of all *PRRT2*-related PKD cases. The most common among the three (82 %) is a frameshift mutation c.649dupC/p.R217Pfs*8 (Méneret *et al.*, 2013). *PRRT2* loss-of-function and haploinsufficiency are suggested as possible causes for pathogenesis due to the truncating nature of the nonsense and frameshift mutations associated with the disease. It is important to better understand the molecular mechanisms leading to *PRRT2*-related PKD for better diagnosis and treatment.

1.4.2. Epidemiology

PKD is a rare disease with an estimated prevalence of 1:150,000 (Chen *et al.*, 2011). There is no unique definition of a rare disease. The Congress of the United States of America defined such a disease as one affecting fewer than 200,000 people in the Orphan Drug Act of 1983. In the European Union such a disease is defined as one affecting 1/2,000 people. PKD, with its prevalence, fulfils the requirements of both classifications. Not many studies have been performed in the area of PKD epidemiology. Palau-Bargues *et al.* (2010) described nine cases in three Spanish families, however neither prevalence nor incidence of PKD in Spain is known. Out of the nine patients, 7 were male. This aligns with the sex ratio of 3-4:1 described in other studies (Singer *et al.*, 2010; Roze *et al.*, 2015).

1.4.3. Paroxysmal dyskinesias

Kinesigenic form of paroxysmal dyskinesia is the most common among the paroxysmal movement disorders. At first, paroxysmal dyskinesias (PxDs) were classified according to the duration of the attack. Later Demirkiran and Jankovic (1995) created a classification of PxDs based on differences between triggers. They recognized four types: non-kinesigenic (PNKD), exercise-induced (PED), hypnogenic (PHD) and PKD. The differences between the PxDs are summarized in table 1.5. PHD is characterized by attacks occurring during slow wave (Non-REM) sleep without having a distinct trigger. In most cases PHD has been subsequently classified as a form of autosomal dominant nocturnal frontal lobe epilepsy (ADNFLE; Erro and Bhatia, 2018).

While mutations in *PRRT2* are the only known cause of PKD, there are many other genes associated with other PxDs. Among them *MR-1*, *KNCMA1* and *ATP1A3* causing PNKD, *SLC2A1*, *ECHS1* and *GCH1* causing PED and *PDHA1*, *PDHX* and *DLAT* which encode for subunits of the pyruvate dehydrogenase complex (PDC) whose deficiency causes PxD meeting the criteria for both PED and PNKD (Erro and Bhatia, 2018).

In contrast with other authors, Erro and Bhatia (2018) suggest a diagnosis of PxDs based on presence of positive signs rather than exclusion. As these positive signs they propose phenomenological variability within-subject with increases in frequency and severity of attacks upon examination, variability of attack duration, presence of several triggers which are non-specific, responsiveness alteration during attacks, neurological and somatic symptoms which cannot be explained medically and irregular response to medications.

Table 1.5. Classification and information summary of paroxysmal dyskinesias according to Demirkiran and Jankovic (1995).

	Paroxysmal kinesigenic dyskinesia (PKD)	Paroxysmal non-kinesigenic dyskinesia (PNKD)	Paroxysmal exercise-induced dyskinesia (PED)	Paroxysmal hypnogenic dyskinesia (PHD)
Trigger	sudden movement	alcohol, coffee, tea, spontaneous	prolonged physical exercise	no triggers (Non-REM sleep stage II)
Duration	seconds to few minutes	a few minutes to 4 hours	5 to 30 minutes	30 to 45 seconds
Typical movement pattern	dystonic, choreatic and ballistic	dystonic	dystonic	dystonic, choreatic and ballistic
Attack frequency	up to 100 per day	2 per day to 3 per year	2 per month to 1 per day	5 per year to 20 per night
Age of onset	7 to 14 years of age	earlier than PKD	9 to 15 years of age	3 to 47 years of age (mean 21.8)
Sex ratio	male>>female	male>female	male>female	male=female
Treatment	anti-epileptic drugs	avoidance of triggers and benzodiazepine	avoidance of triggers and ketogenic diet	anti-epileptic drugs

(data from Schelosky, 2010).

1.4.4. Clinical features and treatment

The term paroxysmal refers to the sudden recurring attacks. Kinesigenic indicates the trigger as being sudden, voluntary movement, acceleration or startle. Dyskinesia refers to involuntary writhing movements which may have a fixed pattern. The paroxysms are therefore characterized by dystonia, chorea (fast dance-like movements) or less commonly

ballismus (flailing movements of the limbs) or a combination of these. The disorder is often misdiagnosed as epilepsy. These, most commonly dystonic attacks occur in extremities and can be preceded by an aura, a crawling or tingling feeling such as limb paresthesia at the site of the paroxysm. The episode lasts up to 5 minutes but most typically only a few seconds, during which the affected person has no alterations of consciousness. If the face or jaw is affected by the episode, speech may be impeded. Any body part can be involved in a unilateral or bilateral attack. Except for the recurring episodes no other symptoms were observed, with electroencephalogram (EEG) and brain imaging studies showing normal results (McGuire *et al.*, 2018).

Typically, the onset of the disorder occurs at 7 to 14 years of age. The symptoms tend to decrease in severity or even resolve completely in adulthood. Although men are affected by the sporadic form of PKD more often than women, there is no such disparity in familial PKD, where the infantile convulsions are more common (McGuire *et al.*, 2018).

The symptoms' improvement shows favourable response to anti-epileptic drugs such as carbamazepine, whose lower doses are sufficient, rather than higher doses used for epilepsy treatment. Other anti-epileptic drugs used for PKD treatment are phenytoin, valproate, oxcarbazepine, lamotrigine, levetiracetam, or topiramate. Whether avoiding triggers such as stress, sleep deprivation or anxiety has any beneficial effect remains uncertain. In patients with BFIE, anti-epileptic drugs and benzodiazepines (psychoactive drugs) may be used for treatment. The adjustment of drug doses is done on individual basis depending on clinical monitoring performed every one or two years (Ebrahimi-Fakhari *et al.*, 2018).

1.4.5. *PRRT2* and paroxysmal kinesigenic dyskinesia

PRRT2 was implicated as the gene causing PKD, ICCA and BFIE by genetic linkage studies (Chen *et al.*, 2011; Schubert *et al.*, 2012; Méneret *et al.*, 2013). Thus far, mutations in *PRRT2* are the only known cause of PKD. The function of *PRRT2* and its role in the pathogenesis of PKD are not fully understood. The *PRRT2* gene has four exons, three of which (2-4) encode a 340 amino acid protein (Figure 1.13). The *PRRT2* protein has an N-terminal domain of 268 amino acid residues with a proline-rich region. Next, a transmembrane helix M1 with a hinge formed by two proline residues occupies amino acids 269-289. This causes M1 not to cross the membrane but fold in half inside the plasma membrane. An intracellular loop occupies residues 290-317. A putative transmembrane

helix M2 (318-338) crosses the membrane leaving an extracellular C-terminal dipeptide occupying residues 339-340 (Rossi *et al.*, 2018; Valtorta *et al.*, 2016). Three isoforms of PRRT2 exist. Isoform 1 being the most common with 340 amino acids. Isoform 2 has 394 residues with an extended C-terminal and isoform 3 has 299 amino acids with the 294-340 replaced by VSPMGP. PRRT2 sequence is highly conserved among mammals (approx. 80 %), and has approximately 30 % similarity with lower vertebrates such as zebrafish (*Danio rerio*; Valtorta *et al.*, 2016).

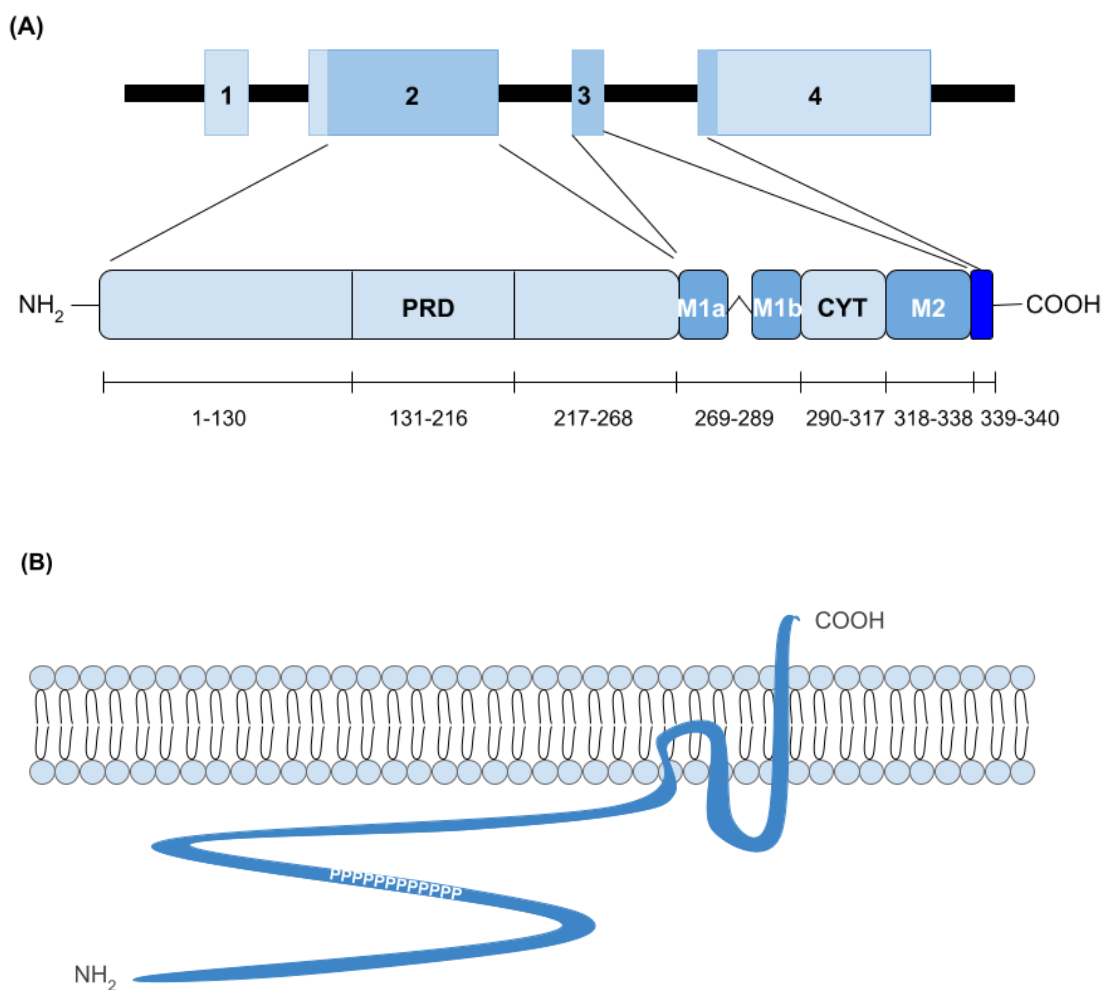


Figure 1.13. (A) Structure of PRRT2 gene and protein. (B) Membrane topology of PRRT2 protein (modified with permission from Valtorta *et al.*, 2016).

PRD – proline-rich domain, M1a and M1b – transmembrane hinged helix, CYT – intracellular loop, M2 – transmembrane helix. See appendix I for license agreement.

Chen *et al.* (2011) identified the CNS as the primary site of *Prprt2* mRNA expression in mice whose protein is highly homologous to the human PRRT2. It localizes at the synaptic membrane putatively interacting with SNAP-25, SYT1, SYT2 and the GluA1 subunit of the AMPA-type glutamate receptor complex (Stelzl *et al.*, 2005; Lee *et al.*, 2012; Schwenk *et al.*, 2014; Li *et al.*, 2015; Valente *et al.*, 2016).

Due to its interaction with the presynaptic proteins involved in neurotransmitter release, PRRT2 is proposed to be involved in synaptic regulation. Valente *et al.* (2016) and Tan *et al.* (2017) suggest a greater involvement of PRRT2 in the presynaptic function. This is considered as one of the pathways by which PRRT2 may be involved in causing PKD, as summarized in table 1.6. It co-distributes with SNAP-25 on the presynaptic membrane and has increased expression during synaptogenesis. In primary neurons, synaptic connections' density is decreased and the ultrastructure of the nerve terminal is changed with *PRRT2* silencing. Also synchronous and spontaneous release of neurotransmitter is impaired without PRRT2. However, Tan *et al.* (2017) do not confirm the SYT1 and SYT2 interaction with PRRT2, as proposed by Valente *et al.* (2016). Most of the *PRRT2* mutations causing PKD are truncating. It was considered that the protein lacking the transmembrane domains is unable to anchor at the presynaptic membrane and gains a detrimental function. Liu *et al.* (2016) showed cytoplasmic location of six different truncating mutant PRRT2. Yet, a complete loss of PRRT2 was observed in brains of mice with a *Prprt2* mutation (Tan *et al.*, 2017) and it was suggested that the nonsense-mediated mRNA decay (NMD) pathway was responsible for the protein's absence (Wu *et al.*, 2014).

Primary excitatory neurons with *PRRT2* knockout have slower exocytosis kinetics, higher neural facilitation and weaker spontaneous and evoked firing. The inhibitory neurons have a stronger basal synaptic transmission and faster depression. Thus, hyperexcitability is proposed as the possible cause of *PRRT2*-associated paroxysmal phenotypes (Valente *et al.*, 2018).

PRRT2 interacts with voltage-gated Na⁺ (Na_v) channels. Sodium ion current of Na_v1.2 and Na_v1.6 channels is decreased by *PRRT2* expression. Haploinsufficiency caused by truncating mutations in *PRRT2* was therefore indicated in hyperactivity of Na_v channels (Fruscione *et al.*, 2018). Neuronal hyperexcitability which may cause paroxysmal attacks may also be the result of PRRT2 affecting glutamate release through its interactions with SNAP-25 and GluA1 (Li *et al.*, 2015).

PRRT2 was also shown to be a negative regulator of synaptic vesicle priming. It blocks vesicle associated v-SNARE assembly with t-SNARE by interaction with SNAP-

25, part of the t-SNARE complex. This impedes opening of the fusion pore and neurotransmitter release. Mutations in *PRRT2* associated with PKD have been shown to interrupt this SNARE regulation (Coleman *et al.*, 2018).

Table 1.6. PRRT2 mechanisms of involvement in PKD pathogenesis.

PRRT2 interaction	PRRT2 function	PRRT2 haploinsufficiency effect
SNAP-25 and GluA1	Regulation of synchronous and spontaneous release of neurotransmitter	Hyperexcitability – dysregulated neurotransmitter release
	Negative regulator of synaptic vesicle priming	
Nav _v 1.2 and Nav _v 1.6 channels	Decreasing sodium ion current	Hyperactivity of Na ⁺ channels

1.4.5.1. *PRRT2* gene

Mutations in *PRRT2* account for ~ 90 % of familial and ~ 30 % of sporadic cases of PKD, depending on the study (Liu *et al.*, 2012). Chen *et al.* (2011) first identified *PRRT2* as the causative gene for PKD, describing three mutations: c.514_517delTCTG (p.S172Rfs*3), c.649dupC (p.R217Pfs*8) and c.972delA (p.V325Sfs*12). Since then there have been over 30 nonsense/frameshift and over 20 missense mutations described (Figure 1.14). Mutation c.649dupC (p.R217Pfs*8) is the most common among them, segregating with approximately 80 % of *PRRT2*-associated familial PKD cases. Novel *PRRT2* mutants affecting PKD patients are still being discovered, often through WES. Spanning a variety of mutations from truncating at the N-terminal c.46G>T (p.E16*; Kita *et al.*, 2017) to missense at the C-terminal c. 959C>T (p.A320V; Lu *et al.*, 2018). Prabhakara and Anbazhagan (2014) have also identified a mutation c.244C>T in 5'-UTR in exon 2 of *PRRT2* in a PKD patient. Idiopathic PKD is more common in men than women. On the other hand, *PRRT2* mutations were identified in patients independently of ethnicity (Liu *et al.*, 2011; Hedera *et al.*, 2012; Lee *et al.*, 2012; Méneret *et al.*, 2012; Ono *et al.*, 2012).

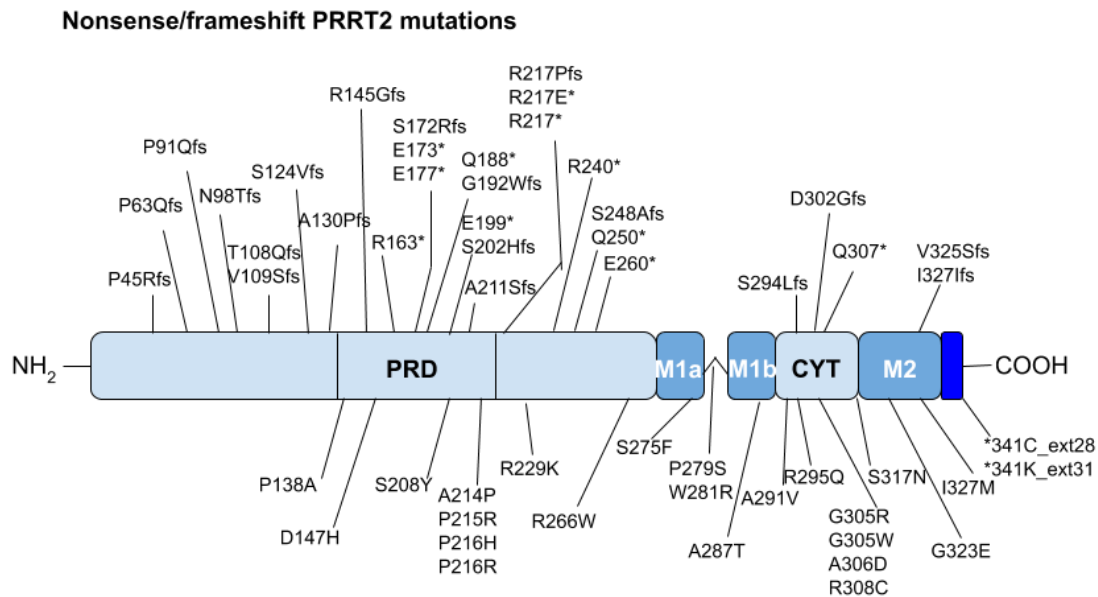


Figure 1.14. Map of nonsense/frameshift and missense mutations in proline-rich transmembrane protein 2 (PRRT2) with regards to their protein domain location (modified with permission from Valtorta *et al.*, 2016).

See appendix I for license agreement.

Mislocalization to the cytoplasm and NMD of truncating mutations of *PRRT2* are meant to be the two processes leading to hyperexcitability in PKD (Wu *et al.*, 2014). Missense mutations in *PRRT2* cluster at the C-terminus unlike the truncating mutations which spread throughout the entire sequence. Interestingly the C-terminal region is the most conserved part of the molecule with up to 90 % similarity among mammals and 60 % in zebrafish (Valtorta *et al.*, 2016). Many *PRRT2* missense mutations often lead to defects in localization to the plasma membrane which may contribute to pathogenesis by the protein's loss-of-function (Tsai *et al.*, 2019).

Mutations in *PRRT2* may also be associated with other phenotypes than PKD. Among them BFIE, ICCA, epilepsy, febrile seizures, infantile non-convulsive seizures (INCS) and nocturnal convulsions (NC; Liu *et al.*, 2011; Liu *et al.*, 2012; Ono *et al.*, 2012).

1.4.6. Other genetic risk factors for paroxysmal kinesigenic dyskinesia

Since over 50 % of patients with primary PKD do not have a mutation in *PRRT2*, a novel *locus* for the disease was mapped to chromosome 3q28-29 (Liu *et al.*, 2016). This

was done using genome-wide short tandem repeats (STR) and SNP based linkage analysis. Alternatively, patients affected by PKD, not harbouring a *PRRT2* mutation were screened for myofibrillogenesis regulator 1 (*MR-1*), solute carrier family 2 member 1 (*SLC2A1*) and chloride voltage-gated channel 1 (*CLCN1*) genes which cause PNKD, PED and myotonia congenita (MC), respectively (Wang *et al.*, 2016). Sixteen out of 28 patients were found to possess variants in these genes.

1.4.7. Nonsense-mediated mRNA decay (NMD) pathway

As mentioned before, Wu *et al.* (2014) have suggested the NMD pathways is involved in decay of the *PRRT2* mRNA with a PTC. Since over thirty nonsense and frameshift mutations in *PRRT2* have been described as the cause of PKD, with mutation c.649dupC (p.R217Pfs*8) as the most common and many other frameshift mutations ending in a PTC, NMD and haploinsufficiency could be essential molecular mechanisms through which PKD is caused.

NMD is a quality control system associated with PTC in open reading frames (ORFs), which eliminates aberrant transcripts. The canonical principle of the NMD pathway is the detection of PTC in transcript, tagging and destruction of the aberrant mRNA. NMD is involved in decay of approximately 10 % of unmutated mRNAs in mammalian cells (Kurosaki *et al.*, 2019), as it also functions as a quantity control system for transcripts. While there are many mechanisms by which transcripts can evade the NMD pathway, reviewed by Dyle *et al.* (2019), and many different substrates for NMD were described in mammalian cells (Schweingruber *et al.*, 2013; Kishor *et al.*, 2019), here I will concentrate on the canonical mechanism of this very conserved (Culbertson *et al.*, 1999) pathway, due to its importance for the study.

PTCs can arise from nonsense or frameshift mutations, as well as DNA rearrangement or erroneous RNA splicing. An unmutated transcript in mammalian cells is left with many RNA binding proteins (RBPs) attached, after splicing. Among them exon-exon junction complexes (EJCs). EJCs are protein complexes deposited upstream of the exon-exon junction, 20-24 nucleotides away (Hug *et al.*, 2016). The translation machinery dislodges RBPs from mRNA and in the case of EJCs the ribosome-associated protein PYM executes this task. As the ribosome continues translation, on its way towards the stop codon, the EJCs are displaced and the protein is forming. However, when a PTC is present, the translation ends prematurely and downstream EJCs remain attached to the transcript. These complexes initiate the NMD pathway and the destruction of the transcript. Thus, nonsense

mutations can lead to loss-of-function and cause diseases through haploinsufficiency. Through many studies it was also established that the PTC should preferably occur > 50-55 nucleotides upstream of an exon-exon junction to trigger NMD response (Dyle *et al.*, 2019).

RNA helicase UPF1 (up-frameshift 1) is an essential part of the NMD machinery which relies on phosphorylation and dephosphorylation events for progression of the transcript destruction. A SMG1c complex (protein kinase SMG1, PI3K-like kinase, SMG8 and SMG9) phosphorylates UPF1. At first, SMG1 associates with UPF1 which in turn interacts with eukaryotic release factors eRF1 and eRF3. Together they form the surveillance complex (SURF). SMG8 and SMG9 regulate SMG1 activity. The SURF complex interacts with UPF2, UPF3b and the downstream EJC in case of a PTC present to form a decay-inducing complex (DECID). This triggers UPF1 phosphorylation and eRF1 and eRF3 dissociate. Through association with UPF2, UPF1 is activated and the complex translocates along the mRNA, downstream of the EJC. It dislodges proteins from mRNA along the way. UPF1 then associates with SMG5, SMG6 and SMG7 and mRNA degradation factors. Endonuclease SMG6 cleaves the transcript and the degradation is initiated (Hug *et al.*, 2016).

2. OBJECTIVES

Alzheimer's disease, primary lateral sclerosis and paroxysmal kinesigenic dyskinesia are three distinct neurological diseases which can be caused by single nucleotide variants in known or unknown disease-causing genes. While the clinical effects of these variants will be dependent on the affected gene, the molecular mechanisms involved in pathogenesis may share common characteristics amongst all three diseases.

The general objective of this thesis is **to comprehend the etiopathogenesis of three neurological diseases affecting three distinct Spanish families** in order to identify the affected molecular pathways in addition to identifying possible therapeutic targets.

The following partial objectives are pursued in order to achieve the general objective:

1. Determine variants possibly responsible for AD, PLS and PKD in the described families.
2. Characterize the role these variants may play in:
 - a. mRNA expression levels,
 - b. protein expression,
 - c. protein stability.
3. Identify possible functional impact of the variants.

3. MATERIALS AND METHODS

3.1. Nomenclature

All gene names used have been approved by the HUGO Gene Nomenclature Committee (HGNC; Braschi *et al.*, 2019; www.genenames.org). The GenBank sequence database (Clark *et al.*, 2016; www.ncbi.nlm.nih.gov/genbank/) was used for variant positions described in the text. The form of describing variants with gene name and the specific amino acid change in superscript is adopted for clarity, due to a notable number of variants discussed.

3.2. Subjects

The DNA samples used in this study were collected from patients belonging to three Spanish families with neurological diseases. Family denoted UGM037 with patients affected by fAD, family UGM471 with patients affected by fPLS and family UGM478 with patients affected by fPKD.

All samples were collected with approval of the corresponding institutional review boards of the corresponding hospitals with signed informed consent from patients.

3.2.1. Family UGM037

DNA samples of eleven individuals were available for this family. Relationships between the individuals are summarized in figure 3.1. Three of the members were diagnosed with fAD: 04-046, 04-048 and 04-049 (assigned these numbers for privacy purposes). Among the analytics used for diagnosis, a computed tomography scan was performed. No additional amyloid positron-emission tomography (PET) nor a study of CSF biomarkers was done.

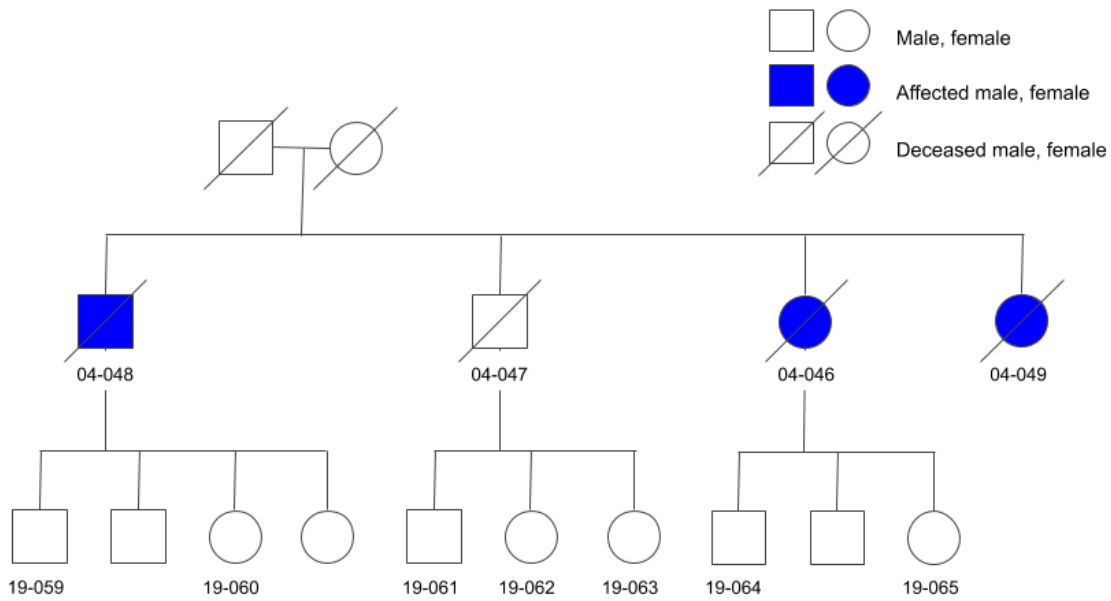


Figure 3.1. Family tree of UGM037 family with the three affected members denoted by coloured in shapes.

Individuals whose DNA samples were available are denoted with an assigned number.

3.2.2. Family UGM471

DNA samples of seven individuals were available for this family. Relationships between the individuals are summarized in figure 3.2. Three members of the family were diagnosed with fPLS, mother (0002) and two of her sons (1001, 1002). No previous family history of the disease was reported.

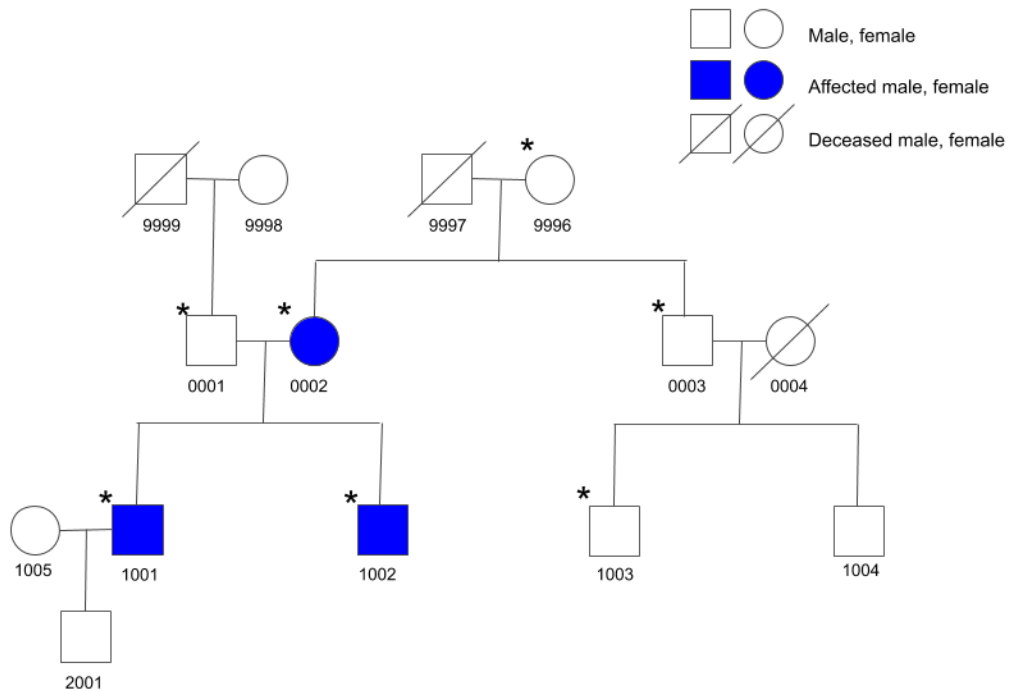


Figure 3.2. Family tree of UGM471 family with the three affected members denoted by coloured in shapes.

Individuals whose DNA samples were available are denoted with an asterisk.

3.2.3. Family UGM478

DNA samples of seven individuals were available for this family. Relationships between the individuals are summarized in figure 3.3. This Spanish family was previously described by Palau-Bargues *et al.* in 2010. Five members of the family were diagnosed with fPKD: 1003, 1005 and their respective children 2004, 2005 and 2007. Patient with the assigned number 2005 was examined using neuroimaging and electroencephalography (EEG). The rest of the patients were diagnosed using clinical history and neurological examination.

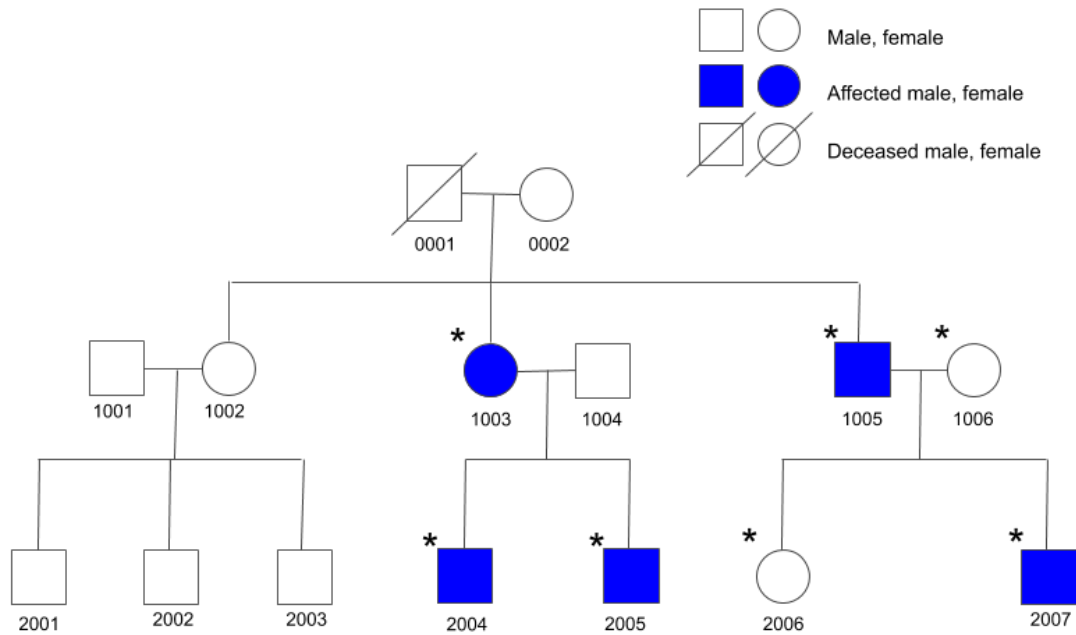


Figure 3.3. Family tree of UGM478 family with the five affected members denoted by coloured in shapes.

Individuals whose DNA samples were available are denoted with an asterisk.

3.2.4. Control population

A control population of healthy individuals was used for variant frequency verification. The control population consisted of 388 samples from healthy people of ages between 18 and 61. 197 of the samples came from men and 191 from women. All of them were of Spanish origin.

Samples from AD patients were also available. 124 patients, among which 97 had probable LOAD, 8 had probable EOAD and 19 had EOAD. 42 among them were men and 82 were women. Just as with the control population, they were of Spanish origin.

3.3. Materials

3.3.1. Bacterial strain

The *Escherichia coli DH5α* strain was used in this work for all manipulations involving the need of electrocompetent bacterial cells. These manipulations are detailed in further sections of this chapter. This *E. coli* strain was developed with specific mutations which allow for its high efficiency transformation. The specific genotype of *E. coli DH5α* is $dlacZ \Delta M15 \Delta (lacZYA-argF) U169 recA1 endA1 hsdR17(rK-mK+) supE44 thi-1 gyrA96 relA1$.

3.3.2. Cell line

SH-SY5Y (ATCC[®] CRL-2266[™]; Figure 3.4) cells were used in all experiments involving the need of a human cell line. SH-SY5Y is a cell line subcloned from SK-N-SH (ATCC[®] HTB-11[™]) parental line derived from a metastatic bone tumour biopsy. Deposited to ATCC[®] by June L. Biedler in 1970, these neuroblastoma cells have an epithelial morphology. SH-SY5Y are often used as *in vitro* models of neuronal function as they can be driven toward adrenergic, cholinergic or dopaminergic phenotypes (Kovalevich and Langford, 2013). As detailed by Biedler *et al.* (1978), the cells have a mixed phenotype of adherent and floating in suspension.

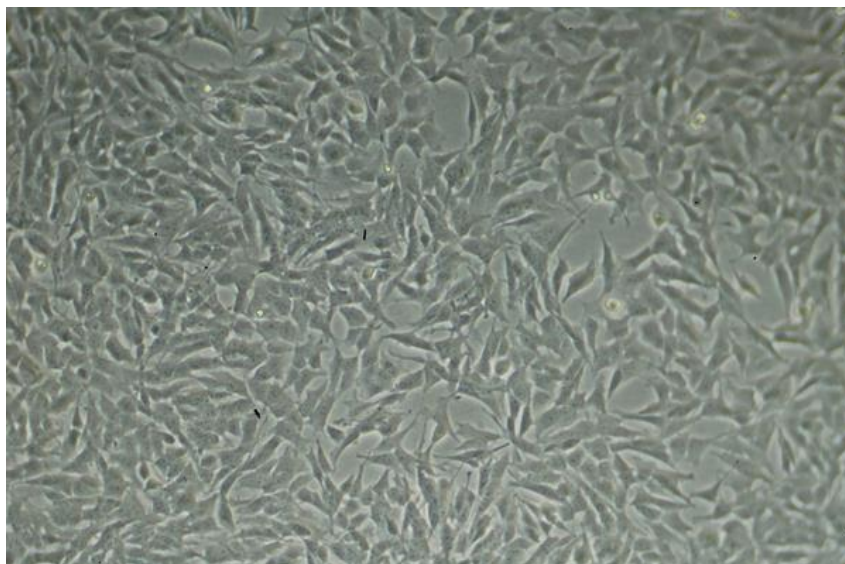


Figure 3.4. SH-SY5Y (ATCC[®] CRL-2266[™]) neuroblastoma cells.

3.3.3. Buffers and media solutions

The recipes for buffers and media used in this work are listed below.

3.3.3.1. Buffers and media used for cultivation of bacteria and genetic engineering

1. Lysogeny broth (LB) medium
 - 1 % (w/v) tryptone (Condalab, Spain)
 - 0.5 % (w/v) yeast extract (Scharlab, Spain)
 - 1 % (w/v) NaCl (Scharlab, Spain)

Dissolve in double-distilled water (ddH₂O). Adjust pH to 7.0 using NaOH and autoclave. Store at 4 °C. Add corresponding antibiotic (ampicillin (1:1000 of 100 mg/ml; AM04680005, Scharlab, Spain), kanamycin (1:1000 of 50 mg/ml; ab60018, Abcam, UK) or chloramphenicol (1:1000 of 33 mg/ml; C0378, Sigma-Aldrich, Merck, Germany)) before use.

2. LB agar plates

1.5% (w/v) agar bacteriological grade (Condalab, Spain) was added to the previously prepared LB medium and autoclaved. The corresponding antibiotic (ampicillin (1:1000 of 100 mg/ml; AM04680005, Scharlab, Spain), kanamycin (1:1000 of 50 mg/ml; ab60018, Abcam, UK) or chloramphenicol (1:1000 of 33 mg/ml; C0378, Sigma-Aldrich, Merck, Germany)) was added to medium after cooling at room temperature to < 50 °C. The medium was poured into Petri dishes (100 x 15 mm; Sarstedt, Germany) and left overnight to solidify.

3. Super optimal broth (SOB)

- 2 % (w/v) tryptone (Condalab, Spain)
- 0.5 % (w/v) yeast extract (Scharlab, Spain)
- 10 mM NaCl (Scharlab, Spain)
- 2.5 mM KCl (Scharlab, Spain)

Dissolve in ddH₂O. Adjust pH to 7.5 using NaOH and autoclave. Store at 4 °C.

4. SOB with catabolite repression (SOC)

After autoclaving SOB, add the following:

- 10 mM MgCl₂ (Scharlab, Spain)
- 10 mM MgSO₄ (Calbiochem, Merck, Germany)
- 20 mM glucose (Scharlab, Spain)

All of the above need to be filtered with a 0.45 µm (Lab Logistics Group GmbH, Germany) filter before use. Store at 4 °C.

5. Resuspension solution for miniprep plasmid isolation

- 25 mM Tris-HCl pH 8.0 (PanReac AppliChem, USA)
- 50 mM glucose (Scharlab, Spain)

- 10 mM Ethylenediaminetetraacetic acid (EDTA; OmniPur[®] EDTA - CAS 60-00-4 – Calbiochem, Merck, Germany)

Dissolve in ddH₂O. Store at 4 °C.

6. Denaturing solution for miniprep plasmid isolation

- 0.2 N NaOH (Sigma-Aldrich, Merck, Germany)
- 1.0 % sodium dodecyl sulfate (SDS) (Scharlab, Spain)

Dissolve in ddH₂O. Store at room temperature.

7. Renaturing solution for miniprep plasmid isolation

- 3 M potassium acetate (Sigma-Aldrich, Merck, Germany)
- 2 M glacial acetic acid (Scharlab, Spain)

Dissolve in ddH₂O. Store at 4 °C.

3.3.3.2. Buffers and media used for protein biochemistry

1. Phosphate buffered saline (PBS) 1X

- 137 mM NaCl (Scharlab, Spain)
- 2.7 mM KCl (Scharlab, Spain)
- 10 mM Na₂HPO₄ (Scharlab, Spain)
- 1.8 mM KH₂PO₄ (Scharlab, Spain)

Dissolve in ddH₂O. Adjust pH to 7.4 with 1M HCl (Scharlab, Spain) and autoclave. Store at 4 °C.

2. Radioimmunoprecipitation assay (RIPA) buffer

- 50 mM Tris-HCl pH 8.0 (PanReac AppliChem, USA)
- 150 mM NaCl (Scharlab, Spain)
- 0.1 % (w/v) SDS (Scharlab, Spain)
- 0.5 % (w/v) deoxycholic acid (Sigma-Aldrich, Merck, Germany)
- 1 % (w/v) Triton X-100 (PanReac AppliChem, USA)
- 1X Protease Inhibitor Cocktail (cOmplete[™], EDTA-free, Roche, Switzerland)

Dissolve in ddH₂O. Store at 4 °C before addition of the protease inhibitor. Add protease inhibitor upon use.

3. Laemmli 5X buffer

- 250 mM Tris-HCl pH 6.8 (PanReac AppliChem, USA)
- 40 % (v/v) glycerol (Scharlab, Spain)
- 4 % (w/v) SDS (Scharlab, Spain)
- 0.005 % (w/v) bromophenol blue (Thermo Fisher Scientific, USA)
- 10 % (v/v) β -mercaptoethanol (Sigma-Aldrich, Merck, Germany)

Dissolve in ddH₂O. Store at -20 °C. Dilute to 1X with protein sample before use.

4. Running buffer 10X for Western Blot

- 250 mM Tris base (PanReac AppliChem, USA)
- 1920 mM glycine (PanReac AppliChem, USA)
- 1 % (w/v) SDS (Scharlab, Spain)

Dissolve in ddH₂O. Store at room temperature. Dilute to 1X before use.

5. Transfer buffer 10X for Western Blot

- 25 mM Tris base (PanReac AppliChem, USA)
- 200 mM glycine (PanReac AppliChem, USA)

Dissolve in ddH₂O. Store at room temperature. Dilute to 1X before use, adding ddH₂O and methanol (Acros Organics, Thermo Fisher Scientific, USA) at 20 % (v/v).

6. TBS Tween 1X

- 20 mM Tris base (PanReac AppliChem, USA)
- 150 mM NaCl (Scharlab, Spain)
- 0.1 % (v/v) Tween 20 (PanReac AppliChem, USA)

Dissolve in ddH₂O. Store at room temperature.

7. Blocking buffer

- TBS Tween 1X
- 10 % (w/v) skim powder milk without calcium

8. Mild stripping buffer

- 200 mM glycine (PanReac AppliChem, USA)
- 0.1 % (w/v) SDS (Scharlab, Spain)
- 1 % (v/v) Tween 20 (PanReac AppliChem, USA)

Dissolve in ddH₂O. Adjust pH to 2.2 with 1M HCl. Store at room temperature.

9. PreScission cleavage buffer (PCB)

- 50 mM Tris base (PanReac AppliChem, USA)
- 150 mM NaCl (Scharlab, Spain)
- 1 mM EDTA (OmniPur[®] EDTA - CAS 60-00-4 – Calbiochem, Merck, Germany)
- 1 mM Dithiothreitol (DTT; DI1360 Scharlab, Spain)

Dissolve in ddH₂O. Adjust pH to 7.5 with 1M HCl. Store at 4 °C.

3.3.3.3. Buffers and media used with SH-SY5Y neuroblastoma cells

1. Complete growth medium

- 88 % MEM Complete Medium with 2 mM L-glutamine (Thermo Fisher Scientific, USA)
- 10 % heat inactivated and filtered fetal bovine serum (FBS; Invitrogen, Thermo Fisher Scientific, USA)
- 1 % non-essential amino acids (NEAAs; Invitrogen, Thermo Fisher Scientific, USA)
- 100 U/ml penicillin-streptomycin antibiotics (cat. number 15140122 Thermo Fisher Scientific, USA)

2. Cryopreservation medium

Supplement the complete growth medium with 5 % (v/v) Dimethyl sulfoxide (DMSO; cat. number D8418, Sigma-Aldrich, Merck, Germany).

3.3.4. Expression vectors

Table 3.1. Expression vectors used in this study.

Empty vectors	Producer
FLAG-HA-pcDNA3.1	Addgene, USA
pCMV-Tag2A	Donation from Jeronimo Bravo laboratory
pFlag – Kan ^R ; C-terminal FLAG epitope	Produced in laboratory
pGEX-6P-1	Addgene, USA
Expression constructs	Producer
<i>ADPRH</i> (BC063883.1) in pGEM-T	Sino Biological, China
<i>ADPRHL1</i> (NM_138430.5) in pcDNA3.1 ⁺ /C-(K)DYK	GenScript, USA
<i>CRI</i> (NM_000651.4) in pcDNA3.1 ⁺ /C-(K)DYK	GenScript, USA
<i>PRRT2</i> (NM_145239.2) in pcDNA3.1 ⁺ /C-(K)DYK	GenScript, USA
pGroEL/GroES	Donation from Anthony Gatenby laboratory

Expression vector maps can be found in appendix II.

All the vectors were verified through sequencing and diagnostic digestion with restriction enzymes and subsequent agarose gel electrophoresis. They were stored in form of glycerol stocks of transformed *E. coli DH5α*. The bacteria (40 μl) were transformed with vectors (50 ng) by electroporation at 1700 V, using the Electroporator 2510 (Eppendorf, Germany). They were incubated at 37 °C in SOC for 45 minutes in a shaking incubator for the cells to generate antibiotic resistance and then plated onto a Petri dish with solid LB with the corresponding antibiotic. The plates were incubated overnight at 37 °C and then single colonies were used for inoculation of liquid LB medium containing the corresponding antibiotic. After incubation at 37 °C overnight, 500 μl of the culture were mixed with 500 μl of 50 % glycerol in a cyrovial (Corning, USA). The cells were stored at -80 °C. Their recovery was done through a miniprep or midiprep plasmid preparation protocol. For miniprep either a QIAprep Spin Miniprep Kit (QIAGEN, Netherlands;

appendix III) was used or a protocol established in the laboratory. For midiprep ZymoPURE II Plasmid Midiprep (Zymo Research, USA; appendix III) kit was used.

For the miniprep protocol established in the laboratory based on Maniatis *et al.* (1982), a sterile autoclaved toothpick was used to scrape the frozen glycerol stock, and the bacteria were streaked onto an LB agar plate containing the corresponding antibiotic. The Petri dish was incubated overnight at 37 °C. A single colony was selected and 3 ml of LB medium containing the corresponding antibiotic were inoculated using a sterile toothpick. The bacterial culture was grown overnight at 37 °C in a shaking incubator. It was then centrifuged at 14,000 x *g* for 4 minutes to pellet the bacteria. The supernatant was discarded and the cells were resuspended in 250 µl of the resuspension solution. Next 250 µl of denaturing solution were added for cell lysis. Finally 250 µl of renaturing solution were added to neutralize the pH for DNA renaturation and SDS precipitation with proteins, lipids and genomic DNA. The precipitate was pelleted by centrifugation at 14,000 x *g* for 10 minutes. The supernatant (approximately 750 µl), with smaller plasmid DNA free in the solution, was transferred into a fresh 1.5 ml tube with 750 µl of chilled isopropanol (Acros Organics, Thermo Fisher Scientific, USA). The DNA was precipitated by centrifugation at 14,000 x *g* for 15 minutes at 4 °C. The supernatant was removed and 70 % ethanol (Acros Organics, Thermo Fisher Scientific, USA) wash was done for excess salt removal. 500 µl of 70 % ethanol were added and the tube was centrifuged at 14,000 x *g* for 5 minutes at 4 °C. After complete removal of ethanol, allowing for the pellet to dry, it was resuspended in 50 µl of Milli-Q[®] ultrapure water (Merck, Germany). The concentration of plasmid DNA was measured by Nanodrop 2000 (Thermo Fisher Scientific, USA). Plasmids were stored at -20 °C.

3.3.5. Antibodies

Table 3.2. Primary and secondary antibodies used in this study.

Antibody	Working concentration	Provider (cat. number)
Primary		
Anti-Actin antibody produced in rabbit	1:1000	Sigma-Aldrich, Merck, Germany (ABT1485)
Anti-ADPRH antibody produced in rabbit (321-335)	1:1000	Sigma-Aldrich, Merck, Germany (A7483)
ANTI-FLAG [®] antibody produced in rabbit	1:1000	Merck, Germany (F7425)
Monoclonal ANTI-FLAG [®] M1 antibody produced in mouse	1:1000	Sigma-Aldrich, Merck, Germany (F3040)
Secondary		
Anti-mouse IgG-HRP antibody produced in goat	1:5000	Invitrogen, Thermo Fisher Scientific, USA (A16084)
Anti-rabbit IgG-HRP antibody produced in goat	1:5000	Santa Cruz Biotechnology, USA (sc2030)
Anti-rabbit IgG-HRP antibody produced in mouse	1:5000	Santa Cruz Biotechnology, USA (sc2357)

3.4. Methods

3.4.1. Whole-exome sequencing

WES is a NGS technique for sequencing the exome of a DNA sample. While GWAS is a good technique for identifying genetic risk factor regions, it is less likely to successfully recognise rare frequency variants responsible for affecting or causing a disease in a particular family. WES is commonly used for identifying rare variants in neurodegenerative diseases (Cukier *et al.*, 2017; Raghavan *et al.*, 2018).

In this study, samples UGM037-04-046 and UGM037-04-048 from the members of the AD affected family and UGM478-1003, UGM478-2004 and UGM478-2005 from the members of the PKD affected family were sequenced using the HiSeq 4000[®] sequencer (Macrogen, Rep. of Korea). Samples UGM471-0002, UGM471-1001 and UGM471-1002

from the members of the PLS affected family were sequenced using the HiSeq 2000[®] sequencer (Sistemas Genomicos, Spain). In both cases WES was performed.

Sample preparation and quality control took place in both systems. Library construction was performed through random fragmentation of DNA and subsequent 5' and 3' adapter ligation. This may be performed in a single tagmentation step. The adapter ligated random fragments were then PCR amplified and gel purified. The library was then loaded onto a flow cell where surface-bound oligonucleotides complementary to the adapters would capture the DNA fragments. The fragments were then amplified by bridge amplification into clonal clusters, thus completing cluster generation. Sequencing can then take place using the reversible terminator-based method developed by Illumina sequencing by synthesis (SBS) which allows for detection of single bases as they incorporate into template strands. Raw data obtained from WES were then analysed through three major steps: base-calling and image analysis, read alignment and SNP calling.

3.4.1.1. Base-calling and image analysis

Samples sequenced using the HiSeq 4000[®] sequencer had images generated by HCS (HiSeq Control Software v3.3). These were then used for system control and base-calling by RTA (Real Time Analysis v2.5.2) software. Illumina package bcl2fastq (v2.16.0.10) was used for conversion of base calls binary into FASTQ files. Phred quality score can be found in the FASTQ files. It assigns the accuracy of each nucleotide a numerical value calculated with $-10\log_{10}P$, where P is the probability of an erroneous base call.

3.4.1.2. Read alignment

All reads were aligned against the human reference genome version GRCh37/hg19 retrieved from the UCSC (University of California Santa Cruz) Genomics Institute; <http://hgdownload.cse.ucsc.edu/goldenPath/hg19/bigZips/chromFa.tar.gz>. Read alignment was performed using BWA (Burrows-Wheeler Alignment; <http://bio-bwa.sourceforge.net/>). The output file was in sai format which was converted to SAM (Sequence Alignment/Map) and further to BAM (Binary Alignment/Map) format using Picard (<http://picard.sourceforge.net/index.shtml>). Further PCR duplicates were marked, local realignment around indels was performed and quality score recalibration was done.

3.4.1.3. SNP calling

SNP calling was done using the Genome Analysis Toolkit Unified Genotyper program (GATK; <https://gatk.broadinstitute.org/hc/en-us>). SNP filtering is performed to filter out the calls of low quality. Next, annotation using annovar (<http://www.openbioinformatics.org/annovar>) and SnpEff (<http://snpeff.sourceforge.net/SnpEff.html>) software was performed to give a final analysis of the possible effect or function of the variants recovered in the process. The annotation databases were: dbSNP version 142 (<http://www.ncbi.nlm.nih.gov/SNP/>), 1000Genome version Phase3 (Auton *et al.*, 2015), ClinVar version 05/2015 (<https://www.ncbi.nlm.nih.gov/clinvar/>) and ESP version ESP6500SI_V2 (<https://evs.gs.washington.edu/EVS/>).

3.4.1.4. Results filtering and identification of interesting variants

From data obtained, the most interesting variants for further investigation were selected through a filtration process. Variants in genes associated with the respective diseases and other neurological disorders were inspected first. Next, variants in other genes were chosen for investigation depending on their role, function, tissue of expression and pathways they are involved in.

These were the primary filters for variant selection:

- Variant present at least in all affected individuals of the family
- Zygosity in accordance with Mendelian inheritance
- Exclusion of synonymous, and previously described as benign, variants
- Sorting Intolerant From Tolerant (SIFT) score prediction considered deleterious (Kumar *et al.*, 2009)
- Polymorphism Phenotyping (Polyphen) score prediction considered damaging (Adzhubei *et al.*, 2010)
- Putative impact prediction considered MODERATE or HIGH by SnpEff

The selected variants were verified using Sanger sequencing. All of the primers used in this work are assigned a number corresponding to a sequence in tables in appendix IV. The variants were assessed using the: “Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology” by Richards *et al.* (2015). Also by the gene constraint observed/expected score developed by Genome

Aggregation Database (gnomAD; Karczewski *et al.*, 2020; <https://gnomad.broadinstitute.org/>) which shows how tolerant a gene is to synonymous, missense or loss-of-function variants. Lower values of the observed/expected score suggest a higher intolerance for a class of variation. A 90 % confidence interval (CI) of under 35 is suggested to be used as a threshold.

3.4.1.5. Public databases

The following public databases were consulted for verification of the variant frequency: Collaborative Spanish Variant Server (CSVS), Trans-Omics for Precision Medicine (TOPMed), Exome Aggregation Consortium (ExAC), gnomAD, NHLBI GO Exome Sequencing Project (GO-ESP) and 1000 Genomes Project (1000G).

3.4.2. Sanger sequencing confirmation of WES results

As suggested by Mu *et al.* (2016), confirmation of NGS variants by Sanger sequencing is necessary for maintaining high sensitivity. For that reason, all the variants selected for further investigation from the WES results were verified by Sanger sequencing.

The PCR conditions used for obtaining amplicons to be used for Sanger sequencing were as follows:

dNTPs 2.5 mM	0.30 μ l
Forward primer (10 μ M)	0.15 μ l
Reverse primer (10 μ M)	0.15 μ l
*Standard Reaction Buffer with MgCl ₂ 10X	1.00 μ l
*Biotools DNA Polymerase (1 U/ μ l)	0.30 μ l
DNA (50 ng/ μ l)	0.30 μ l
H ₂ O	7.80 μ l

Primers can be found in appendix IV, table A1.

*Producer: Biotools, Spain

The programs used were based on the following PCR steps:

Step	Number of cycles	Temperature	Time
Initial denaturation	1	95 °C	10 min
Denaturation		95 °C	30 sec
Annealing	35 – 40*	55 – 65 °C*	30 sec
Extension		72 °C	45 sec
Final extension	1	72 °C	10 min
Cooling	1	4 °C	∞

*Depending on the primers used, the number of cycles and the annealing temperature varied for optimal reaction specificity and sensitivity.

After the PCR, 5 µl of the products were incubated with 2 µl of ExoSAP-IT™ PCR Product Cleanup Reagent (cat. number 78200.200.UL, Thermo Fisher Scientific, USA) at 37 °C for 15 minutes and at 80 °C for another 15 minutes. Sanger sequencing was performed by the DNA analysis service at the Institute of Biomedicine of Valencia (Servei d'anàlisi d'ADN del IBV) using one of the pair of primers used for amplicon formation at a concentration of 3.2 µM.

3.4.3. Allele specific PCR

Allele-specific PCR (ASPCR) was used to study variant frequency in control population for each single nucleotide variant (SNV) selected for further study. ASPCR allows for detection of point mutations in DNA by synthesis of an oligonucleotide primer with 3' mismatch with the DNA template (Ugozzoli and Wallace, 1991). Using forward and reverse oligonucleotide primers homologous to the DNA template and others where one primer has a 3' mismatch with the DNA template, both in separate reactions, will give different results under optimized conditions. Primers used for ASPCR are listed in table A2 in appendix IV.

The PCR conditions used for ASPCR were as follows:

dNTPs 2.5 mM	0.30 μ l
Forward primer (10 μ M)	0.15 μ l
Reverse primer (10 μ M)	0.15 μ l
*Standard Reaction Buffer with MgCl ₂ 10X	1.00 μ l
*Biotools DNA Polymerase (1 U/ μ l)	0.30 μ l
DNA (50 ng/ μ l)	0.30 μ l
H ₂ O	7.80 μ l

Primers can be found in appendix IV, table A2.

*Producer: Biotools, Spain

The programs used were based on the following PCR steps:

Step	Number of cycles	Temperature	Time
Initial denaturation	1	95 °C	10 min
Denaturation		95 °C	30 sec
Annealing	35 – 40*	55 – 65 °C*	30 sec
Extension		72 °C	45 sec
Final extension	1	72 °C	10 min
Cooling	1	4 °C	∞

*Depending on the primers used, the number of cycles and the annealing temperature varied for optimal reaction specificity and sensitivity.

The conditions were always optimized to clearly see the bands for well amplified products. This would allow for distinguishing between bands for primers homologous with the DNA template, and for primers where one had a 3' mismatch with the DNA template.

3.4.4. Plasmid manipulations

3.4.4.1. Site-directed mutagenesis

Site-directed mutagenesis was used to introduce specific single nucleotide variants into cDNAs of *CRI* and *ADPRH*. For *CRI* a FLAG epitope followed by a stop codon was introduced at the appropriate site. For *ADPRH* four distinct variants were introduced.

A PCR was performed to amplify the plasmid with a relevant change, introduced by appropriate design of the primers. The PCR conditions used for site-directed mutagenesis using the *CRI* (NM_000651.4) in pcDNA3.1⁺/C-(K)DYK plasmid were as follows:

dNTPs 2.5 mM	3.0 µl
Forward primer 304f (10 µM)	1.5 µl
Reverse primer 304r (10 µM)	1.5 µl
*Cloned <i>Pfu</i> DNA polymerase reaction buffer 10X	5.0 µl
* <i>Pfu turbo</i> Polymerase (2.5 U/µl)	1.0 µl
DMSO	2.5 µl
DNA (50 ng/µl)	1.5 µl
H ₂ O	34.0 µl

Primers can be found in appendix IV, table A3.

*Producer: Agilent, USA

The programs used were based on the following PCR steps:

Step	Number of cycles	Temperature	Time
Initial denaturation	1	92 °C	2 min
Denaturation		92 °C	30 sec
Annealing	3	50 °C	30 sec
Extension		68 °C	15 min
Denaturation		92 °C	30 sec
Annealing/Extension	35	68 °C	15 min
Final extension	1	72 °C	10 min
Cooling	1	4 °C	∞

The PCR conditions used for site-directed mutagenesis using the *ADPRH* (BC063883.1) in pGEM-T plasmid were as follows:

dNTPs 2.5 mM	2 μ l
Forward primer 300f – 303f (10 μ M)	1 μ l
Reverse primer 300r – 303r (10 μ M)	1 μ l
*Cloned <i>Pfu</i> DNA polymerase reaction buffer 10X	5 μ l
* <i>Pfu turbo</i> Polymerase (2.5 U/ μ l)	1 μ l
DNA (50 ng/ μ l)	1 μ l
H ₂ O	39 μ l

Primers can be found in appendix IV, table A3.

*Producer: Agilent, USA

The programs used were based on the following PCR steps:

Step	Number of cycles	Temperature	Time
Initial denaturation	1	95 °C	2 min
Denaturation		95 °C	30 sec
Annealing	35	52 °C	30 sec
Extension		72 °C	6 min
Final extension	1	72 °C	10 min
Cooling	1	4 °C	∞

After the corresponding PCR was finished, the products were treated with 1 μ l of *DpnI* FastDigest (Thermo Fisher Scientific, USA) endonuclease in 6 μ l of FastDigest buffer 10X, filled until 60 μ l with Milli-Q[®] ultrapure water. Incubation took place for 30 minutes at 37 °C. *DpnI* recognises the sequence 5'-G^{m6}ATC-3', cutting after the adenine base only when the recognition site is methylated. This allows for degradation of the template, leaving the newly formed plasmid intact. Further sample purification was performed using the High Pure PCR Cleanup Micro Kit (Roche, Switzerland; appendix III). The product was ready for *E. coli DH5a* transformation by electroporation as described in the “Expression vectors” section (3.3.4.). Single colonies were grown on a Petri dish with LB and the corresponding antibiotic. The plasmids were extracted by miniprep procedure and the success of mutagenesis was verified by Sanger sequencing (as described in section 3.4.2.).

3.4.4.2. Subcloning

In order to avoid artefacts introduced by PCR employed in the mutagenesis, the mutated and verified insert was subcloned into a fresh vector. In this way, I ensure that no mutation affecting, for example, antibiotic resistance or promoter, is transferred to subsequent experiments. *CRI* (NM_000651.4) in pcDNA3.1⁺/C-(K)DYK and its variant introduced through site-directed mutagenesis, both with a sequence for the FLAG epitope, were subcloned into a pcDNA3.1⁺/C-(K)DYK vector. *ADPRH* (BC063883.1) in pGEM-T and its variants introduced through site-directed mutagenesis were subcloned into a pFlag vector with kanamycin resistance cassette and a C-terminal FLAG epitope sequence; and also into pGEX-6P-1 vector. *PRRT2* (NM_145239.2) in pcDNA3.1⁺/C-(K)DYK was subcloned into pCMV-Tag2A vector with an N-terminal FLAG epitope sequence together with three of its truncated variants. All the above mentioned vectors can be found in table 3.1. The following subcloning conditions and programs were used:

The PCR conditions used for subcloning *CRI* (NM_000651.4) in pcDNA3.1⁺/C-(K)DYK and its variant into pcDNA3.1⁺/C-(K)DYK vector were as follows:

dNTPs 2.5 mM	2 µl
Forward primer 400f, 401f (10 µM)	1 µl
Reverse primer 400r, 401r (10 µM)	1 µl
*Cloned <i>Pfu</i> DNA polymerase reaction buffer 10X	5 µl
* <i>Pfu turbo</i> Polymerase (2.5 U/µl)	1 µl
DNA (50 ng/µl)	1 µl
H ₂ O	39 µl

Primers can be found in appendix IV, table A4.

*Producer: Agilent, USA

The programs used were based on the following PCR steps:

Step	Number of cycles	Temperature	Time
Initial denaturation	1	95 °C	2 min
Denaturation		95 °C	30 sec
Annealing	35	60 °C	30 sec
Extension		68 °C	7 min 30 sec
Final extension	1	72 °C	10 min
Cooling	1	4 °C	∞

The PCR conditions used for subcloning *ADPRH* (BC063883.1) in pGEM-T and its variants into pFlag and pGEX-6P-1 vectors were as follows:

dNTPs 2.5 mM	2 µl
Forward primer 402f, 403f (10 µM)	1 µl
Reverse primer 402r, 403r (10 µM)	1 µl
*Cloned <i>Pfu</i> DNA polymerase reaction buffer 10X	5 µl
* <i>Pfu turbo</i> Polymerase (2.5 U/µl)	1 µl
DNA (50 ng/µl)	1 µl
H ₂ O	39 µl

Primers can be found in appendix IV, table A4.

*Producer: Agilent, USA

The programs used were based on the following PCR steps:

Step	Number of cycles	Temperature	Time
Initial denaturation	1	95 °C	2 min
Denaturation		95 °C	30 sec
Annealing	35	60 °C	30 sec
Extension		72 °C	1 min
Final extension	1	72 °C	10 min
Cooling	1	4 °C	∞

The PCR conditions used for subcloning *PRRT2* (NM_145239.2) in pcDNA3.1⁺/C-(K)DYK and its variants into pCMV-Tag2A vector were as follows:

dNTPs 2.5 mM	2 μ l
Forward primer 404f – 407f (10 μ M)	1 μ l
Reverse primer 404r – 407r (10 μ M)	1 μ l
*Cloned <i>Pfu</i> DNA polymerase reaction buffer 10X	5 μ l
* <i>Pfu turbo</i> Polymerase (2.5 U/ μ l)	1 μ l
DNA (50 ng/ μ l)	1 μ l
H ₂ O	39 μ l

Primers can be found in appendix IV, table A4.

*Producer: Agilent, USA

The programs used were based on the following PCR steps:

Step	Number of cycles	Temperature	Time
Initial denaturation	1	95 °C	2 min
Denaturation		95 °C	30 sec
Annealing	35	60 °C	30 sec
Extension		68 °C	1 min
Final extension	1	72 °C	10 min
Cooling	1	4 °C	∞

After the corresponding PCR was finished, the products and the respective vector plasmid were treated with the corresponding FastDigest restriction enzymes (Thermo Fisher Scientific, USA; Table A4, appendix IV) in FastDigest buffer 10X for 30 minutes at 37 °C. The restriction enzyme treated amplicons and vectors were loaded onto a 0.8 % agarose gel with peqGREEN dye (0,004 %; Peqlab, VWR, Germany) for staining of dsDNA, ssDNA and RNA. The molecular weight ladder used in all agarose gels in this study was 1 Kb Plus DNA Ladder (cat. number 11573127, Invitrogen, Thermo Fisher Scientific, USA). The bands were then excised from the gel and purified using the MinElute Gel Extraction Kit (QIAGEN, Netherlands; appendix III). Dephosphorylation of the vector and ligation was performed with the Rapid DNA Dephos & Ligation Kit (Roche,

Switzerland; appendix III). After purification of the sample with the High Pure PCR Cleanup Micro Kit (Roche, Switzerland; appendix III) it was ready for *E. coli DH5α* transformation by electroporation as described in the “Expression vectors” section (3.3.4.). Single colonies were grown on a Petri dish with LB and the corresponding antibiotic. The plasmids were extracted by miniprep procedure and the success of subcloning was verified by Sanger sequencing (as described in section 3.4.2.).

3.4.4.3. Plasmid constructs

The site-directed mutagenesis and subcloning allowed for construction of plasmids necessary for this study (Table 3.3).

Table 3.3. Plasmid constructs created through site-directed mutagenesis and subcloning.

Expression constructs	Plasmid construct produced in the laboratory
<i>ADPRH</i> (BC063883.1) in pGEM-T	<i>ADPRH</i> (BC063883.1) in pFlag
	<i>ADPRH</i> _p.R295P in pFlag
	<i>ADPRH</i> _p.R295Q in pFlag
	<i>ADPRH</i> _p.R295* in pFlag
	<i>ADPRH</i> _p.D55A_D56A in pFlag
	<i>ADPRH</i> (BC063883.1) in pGEX-6P-1
	<i>ADPRH</i> _p.R295P in pGEX-6P-1
	<i>ADPRH</i> _p.D55A_D56A in pGEX-6P-1
<i>CRI</i> (NM_000651.4) in pcDNA3.1 ⁺ /C-(K)DYK	<i>CRI</i> _p.R136* in pcDNA3.1 ⁺ /C-(K)DYK
<i>PRRT2</i> (NM_145239.2) in pcDNA3.1 ⁺ /C-(K)DYK	<i>PRRT2</i> (NM_145239.2) in pCMV-Tag2A
	<i>PRRT2</i> _p.Q106* in pCMV-Tag2A
	<i>PRRT2</i> _p.Q163* in pCMV-Tag2A
	<i>PRRT2</i> _p.Q250* in pCMV-Tag2A

3.4.5. *In silico* analysis

Structural effects of variants were analysed using HOPE web service developed at the Centre for Molecular and Biomolecular Informatics CMBI at Radboud University in Nijmegen (Venselaar *et al.*, 2010; <https://www3.cmbi.umcn.nl/hope/>). RCSB Protein Data

Bank (PDB; Berman *et al.*, 2000; <https://www.rcsb.org/>) was used for recovery of the protein structures deposited by other authors. Prediction of protein stability upon amino acid changes was done using the I-Mutant 2.0 web server (Capriotti *et al.*, 2005; <http://gpcr2.biocomp.unibo.it/cgi/predictors/I-Mutant3.0/I-Mutant3.0.cgi>). Evolutionary conservation of amino acid positions was estimated using ConSurf server (Ashkenazy *et al.*, 2016; <https://consurf.tau.ac.il/>), a bioinformatics web tool. A multiple sequence alignment with hierarchical clustering was done using MultAlin (Corpet, 1988; <http://multalin.toulouse.inra.fr/multalin/>).

3.4.6. Genotyping

3.4.6.1. Genotyping apolipoprotein E gene polymorphism

Allele $\epsilon 4$ of *APOE* is the main genetic risk factor for AD. Genotyping *APOE* polymorphism is therefore important in assessing the cause of AD in patients. The genotyping was done using a PCR followed by a double enzymatic digestion, agarose gel electrophoresis and results interpretation. The PCR was performed under the following conditions:

dNTPs 2.5 mM	1.1 μ l
Forward primer 600f (10 μ M)	0.55 μ l
Reverse primer 600r (10 μ M)	0.55 μ l
DMSO	2.5 μ l
*Standard Reaction Buffer with MgCl ₂ 10X	5 μ l
*Biotools DNA Polymerase (1 U/ μ l)	0.8 μ l
DNA (50 ng/ μ l)	3 μ l
H ₂ O	36.5 μ l

Primers can be found in appendix IV, table A6.

*Producer: Biotools, Spain

The PCR program selected was the following:

Step	Number of cycles	Temperature	Time
Initial denaturation	1	95 °C	10 min
Denaturation		95 °C	30 sec
Annealing	5	60 °C	30 sec
Extension		72 °C	45 sec
Denaturation		95 °C	30 sec
Annealing	20	60 °C with $\Delta T = -0.5$ °C/cycle	30 sec
Extension		72 °C	45 sec
Denaturation		95 °C	30 sec
Annealing	10	50 °C	30 sec
Extension		72 °C	45 sec
Final extension	1	72 °C	5 min
Cooling	1	4 °C	∞

5 μ l of the PCR products were loaded onto a 1.5 % agarose gel with 2 μ l of loading buffer, 3 μ l of molecular weight ladder and run at 120 V for approximately 40 minutes. If the PCR was successful and the bands were visible on the gel, the enzymatic digestion of the samples took place. The following mixture was prepared:

PCR product	13.1 μ l
* <i>Afl</i> III restriction enzyme	0.2 μ l
* <i>Hae</i> II restriction enzyme	0.2 μ l
*Buffer 10X	1.5 μ l

*Producer: New England Biolabs, USA

The mix of the restriction enzymes with the PCR products was then incubated overnight at 37 °C. The digestion products (15 μ l) were loaded onto a 2.5 % agarose gel with 3 μ l of loading buffer, 4 μ l of molecular weight ladder and run at 120 V for approximately 40 minutes. The results interpretation was done according to the following table 3.4.

Table 3.4. Interpretation of the *APOE* gene polymorphism genotyping.

Genotype	Band size (bp)
$\epsilon 2/\epsilon 2$	205
$\epsilon 2/\epsilon 3$	205 + 141
$\epsilon 2/\epsilon 4$	205 + 169
$\epsilon 3/\epsilon 3$	141
$\epsilon 3/\epsilon 4$	141 + 169
$\epsilon 4/\epsilon 4$	169

3.4.6.2. Genotyping *CR1* gene polymorphism

Members of the family affected by AD had their DNA samples genotyped for the rs3818361 SNP and the long isoform of *CR1* ($CR1^*2$), which contribute to AD susceptibility (Mahmoudi *et al.*, 2015; Li *et al.*, 2015). The genotyping of rs3818361 was done using a PCR followed by Sanger sequencing. The PCR for rs3818361 genotyping was performed under the following conditions:

dNTPs 2.5 mM	0.30 μ l
Forward primer 601f (10 μ M)	0.15 μ l
Reverse primer 601r (10 μ M)	0.15 μ l
*Standard Reaction Buffer with $MgCl_2$ 10X	1.00 μ l
*Biotools DNA Polymerase (1 U/ μ l)	0.30 μ l
DNA (50 ng/ μ l)	0.30 μ l
H ₂ O	7.80 μ l

Primers can be found in appendix IV, table A6.

*Producer: Biotools, Spain

The PCR program selected was the following:

Step	Number of cycles	Temperature	Time
Initial denaturation	1	95 °C	10 min
Denaturation		95 °C	30 sec
Annealing	5	60 °C	30 sec
Extension		72 °C	45 sec
Denaturation		95 °C	30 sec
Annealing	20	60 °C with $\Delta T = -0.5$ °C/cycle	30 sec
Extension		72 °C	45 sec
Denaturation		95 °C	30 sec
Annealing	10	50 °C	30 sec
Extension		72 °C	45 sec
Final extension	1	72 °C	10 min
Cooling	1	4	∞

The genotyping of *CRI*2* long isoform was done using a PCR followed by loading of the products onto a 1.5 % agarose gel with 2 μ l of loading buffer, 3 μ l of molecular weight ladder and run at 120 V for approximately 40 minutes. The design of the primers allowed for the amplicons to form only if the genomic DNA had the long isoform of *CRI* (Kucukkilic *et al.*, 2018). The PCR for *CRI*2* long isoform genotyping was performed under the following conditions:

dNTPs 2.5 mM	0.30 μ l
Forward primer 602f (10 μ M)	0.15 μ l
Reverse primer 602r (10 μ M)	0.15 μ l
*Standard Reaction Buffer with MgCl ₂ 10X	1.00 μ l
*Biotools DNA Polymerase (1 U/ μ l)	0.30 μ l
DNA (50 ng/ μ l)	0.30 μ l
H ₂ O	7.80 μ l

Primers can be found in appendix IV, table A6.

*Producer: Biotools, Spain

The PCR program selected was the following:

Step	Number of cycles	Temperature	Time
Initial denaturation	1	95 °C	2 min
Denaturation		95 °C	30 sec
Annealing	20	70 °C with $\Delta T = -0.5$ °C/cycle	30 sec
Extension		70 °C	30 sec
Denaturation		95 °C	30 sec
Annealing	15	60 °C	30 sec
Extension		70 °C	30 sec
Final extension	1	70 °C	5 min
Cooling	1	4 °C	∞

3.4.7. SH-SY5Y cell culture maintenance and transfection

3.4.7.1. Maintenance

SH-SY5Y cells were stored in cryovials (Corning, USA) in liquid nitrogen. Approximately 2.8×10^6 per cryovial. They were submerged in 1 ml of the cryopreservation medium (section 3.3.3.3.). After thawing, the cells were suspended in 9 ml of the complete growth medium (section 3.3.3.3.). Next steps were:

- centrifugation at $125 \times g$ for 3 minutes,
- aspiration of the medium to remove any DMSO present in the cryopreservation medium,
- resuspension in the complete growth medium and plating.

First the cells were grown in a 25 cm^2 cell culture flask, and later transferred onto 35 mm culture dishes (Table 3.5). The medium was exchanged approximately every 48 hours. The flask or plates with the cells in medium were kept at 37 °C in 5 % CO₂. For passage steps, at 50 – 60 % confluence:

- the medium was aspirated,
- trypsin-EDTA (0.05 % with phenol red, cat. number 25300054, Thermo Fisher Scientific, USA) was added for 1 minute at room temperature,
- trypsin-EDTA was then aspirated and the cells were incubated at 37 °C in 5 % CO₂ for 5 minutes,

- the cells were then suspended in the complete growth medium,
- centrifugation at 125 x g for 3 minutes took place,
- aspiration of the medium to remove any trypsin-EDTA residue,
- resuspension in complete growth medium and plating.

For all experimental assays, the number of cells used was estimated at 1.2×10^6 ; grown on the 35 mm culture dish.

Table 3.5. Flasks and dish used for SH-SY5Y cell culture.

Material	Surface area (cm ²)	Number of cells at confluence	Provider
35 mm culture dish	8.8	1.2×10^6	SPL Life Sciences, Rep. of Korea
T-25 flask	25	2.8×10^6	Thermo Fisher Scientific, USA
T-175 flask	175	23.3×10^6	Thermo Fisher Scientific, USA

3.4.7.2. Transfection

At confluence of 70 – 90 % the SH-SY5Y cells were transfected using the Lipofectamine 2000 Transfection Reagent (Thermo Fisher Scientific, USA). Lipofection protocol, previously optimised at the laboratory, was as follows:

1. Dilute 10 µl of lipofectamine in 240 µl of Opti-MEM medium (cat. number 31985062, Thermo Fisher Scientific, USA) per 35 mm plate.
2. Dilute 2500 ng of plasmid DNA with Opti-MEM medium to a final volume of 250 µl.
3. Mix 250 µl of diluted lipofectamine and 250 µl of diluted plasmid DNA and incubate at room temperature for 15 minutes.
4. Add 500 µl of lipofectamine-plasmid DNA complex per plate.
5. Incubate the cells at 37 °C in 5 % CO₂ for 5-6 hours.
6. Add complete growth medium to the cells and incubate at 37 °C in 5 % CO₂ for 48 hours, then analyse the cells.

3.4.8. Nonsense-mediated mRNA decay (NMD) pathway inhibition analysis

NMD inhibition was performed using NMDI14 (cat. number SML1538, Sigma-Aldrich, Merck, Germany), an inhibitor of the NMD pathway which disrupts interactions between SMG7 and UPF1. These proteins are essential components of the mRNA decay complex which targets transcripts with PTCs for elimination (Martin *et al.*, 2014).

For the purpose of NMD inhibition, two groups of cells were prepared. First group consisted of 35 mm culture dishes of SH-SY5Y cells transfected with corresponding plasmids. The second group was identical to the first except after incubation of the cells with the lipofectamine-plasmid DNA complex, NMDI14 at 50 μ M was added. Since NMDI14 was dissolved in 0.2 mg/ml DMSO, the same amount of DMSO dissolved in water was added to the first group of dishes. Except for plasmids containing the constructs studied, an empty vector plasmid and a mock, without plasmid, were used as controls in both groups.

The groups treated with NMDI14 for NMD pathway inhibition are summarised in table 3.6. Each group was duplicated and one duplicate was treated with NMDI14.

Table 3.6. Groups of SH-SY5Y cells transfected with corresponding plasmids used in NMD pathway inhibition analysis.

Group used for Alzheimer's disease related study (SH-SY5Y cells transfected with these plasmids)	
Plasmid	Purpose
<i>CRI</i> (NM_000651.4) in pcDNA3.1+/C-(K)DYK	Wild-type protein expression
<i>CRI</i> _p.R136* in pcDNA3.1+/C-(K)DYK	Truncated protein expression (<i>CR1</i> ^{R136*})
FLAG-HA-pcDNA3.1	Negative control
No plasmid used	Mock control
Group used for paroxysmal kinesigenic dyskinesia disease related study (SH-SY5Y cells transfected with these plasmids)	
Plasmid	Purpose
<i>PRRT2</i> (NM_145239.2) in pCMV-Tag2A	Wild-type protein expression
<i>PRRT2</i> _p.Q106* in pCMV-Tag2A	Truncated protein expression (<i>PRRT2</i> ^{Q106*})
<i>PRRT2</i> _p.Q163* in pCMV-Tag2A	Truncated protein expression (<i>PRRT2</i> ^{Q163*})

Table 3.6. Groups of SH-SY5Y cells transfected with corresponding plasmids used in NMD pathway inhibition analysis (continued).

<i>PRRT2</i> _p.Q250* in pCMV-Tag2A	Truncated protein expression (<i>PRRT2</i> ^{Q250*})
pCMV-Tag2A	Negative control
No plasmid used	Mock control

3.4.9. mRNA levels analysis

3.4.9.1. Cellular RNA extraction and quantification

Extraction of cellular RNA was performed 48 hours after transfection. The QIAamp RNA Blood Mini Kit (QIAGEN, Netherlands; appendix III) was used, following the producer's protocol for extraction of RNA from monolayer cell culture. The extracted RNA was stored at -80 °C after quantification using a Nanodrop 2000 (Thermo Fisher Scientific, USA).

3.4.9.2. DNase I treatment and reverse transcription

Treatment of the extracted cellular RNA with DNase I (cat. number 04716728001, Roche, Switzerland) was required for degradation of any residual DNA. RNaseOUT™ Recombinant Ribonuclease Inhibitor (Invitrogen, Thermo Fisher Scientific, USA) was used to prevent RNA degradation throughout the process. Total RNAs extracted were diluted in nuclease-free water to equalize their concentrations for reverse transcription. The following reaction mixture was prepared for DNase I treatment:

RNaseOUT™	0.25 µl
DNase I (10 U/µl)	1.10 µl
Buffer 10X	0.90 µl
RNA (1 µg) + nuclease-free water	6.75 µl

The reaction took place at 37 °C for 30 minutes and the DNase I inactivation at 72 °C for 10 minutes. Random hexamers at 6 µM were added to the 9 µl mixture for a total volume of 10.5 µl. Annealing took place at 65 °C for 10 minutes. Expand Reverse Transcriptase (Roche, Switzerland; appendix III) kit was used for reverse transcription. The following mixture was prepared:

RNaseOUT™	0.5 µl
dNTPs 10 mM	2.0 µl
DTT 100 mM	2.0 µl
Buffer 5X	4.0 µl
Reverse transcriptase (50 U/ µl)	1.0 µl

The mixture was added to the previous 10.5 µl and incubated at 30 °C for 10 minutes and then at 42 °C for 1 hour. The synthesised cDNAs were quantified using Nanodrop 2000 and stored at -20 °C.

3.4.9.3. Quantitative PCR (qPCR)

For the second step of the reverse transcription-qPCR (RT-qPCR) we used a Fast 7500 Real Time PCR System (Applied Biosystems, Thermo Fisher Scientific, USA). Oligonucleotide primers used for qPCR are listed in table A5 appendix IV. Reverse transcribed cDNAs, from 1 µg of RNA each, were used as the template. Each reaction consisted of:

*SYBR™ Green PCR Master Mix	7.5 µl
Forward primer 500f – 503f (10 µM)	1.0 µl
Reverse primer 500r – 503r (10 µM)	1.0 µl
cDNA	1.0 µl
H ₂ O	4.5 µl

Primers can be found in appendix IV, table A5.

*Producer: Thermo Fisher Scientific, USA

Each sample tested had three technical and two biological replicates. The data obtained was analysed using the LinRegPCR Software developed by Dr Jan Ruijter (Ramakers *et al.*, 2003). All samples were normalised using qPCR results for neomycin sequence amplification, which was encoded in the vectors used. I use the LinRegPCR Software instead of the delta-delta Ct method (also $2^{-\Delta\Delta Ct}$ method), commonly used for calculating the relative fold gene expression from samples obtained from qPCR, because it takes into account the efficiency of the reaction. Taking the exponential base of 2 for the $2^{-\Delta\Delta Ct}$ method assumes a 100 % efficiency which has been proven to not be accurate (Tuomi

et al., 2010). Also the method applied in the LinRegPCR Software calculates efficiency value per gene rather than per reaction which has been shown to give more reliable results (Ruijter *et al.*, 2013).

3.4.10. Protein expression analysis

3.4.10.1. Protein extraction and quantification

Extraction of proteins from plasmid and mock transfected SH-SY5Y cells was performed 48 hours after transfection. Every experiment had three biological replicates. After aspiration of the medium, the cells were washed in cold PBS 1X. After aspiration of PBS 1X, 250 µl of RIPA was added to each 35 mm culture dish with a cell culture monolayer. After an incubation of 5 minutes on ice, the cells were scraped and collected into 1.5 ml microcentrifuge tubes. To increase the efficiency of extraction, sonication was performed using the Bioruptor (Diagenode, Belgium) with 10 cycles of 30 seconds pulse and 30 seconds pause in-between. After sonication the samples were centrifuged at 14,000 x g for 15 minutes at 4 °C. Both the supernatant (used for further study) and the pellet were stored at -20 °C.

Quantification of the protein samples was performed using either the Qubit fluorometer (Invitrogen, Thermo Fisher Scientific, USA).

3.4.10.2. Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE)

In order to separate the proteins according to their molecular mass the SDS-PAGE procedure was employed. Polyacrylamide gels of 10-14 % were prepared. The protein samples were diluted in H₂O according to their concentrations, to load the same amount of each sample onto the gel. Laemmli 5X buffer was added to the samples (to final concentration of 1X) and incubation of 10 minutes at 100 °C for protein denaturation followed. The samples were loaded onto the gel with the Precision Plus Protein™ Dual Color Standards (Biorad, USA) or Spectra™ Multicolor Broad Range Protein Ladder (Thermo Fisher Scientific, USA). The vertical gel was run in running buffer for approximately 2 hours at 100 V.

3.4.10.3. Western Blot (WB)

After separation of the proteins according to their molecular mass with SDS-PAGE they were transferred onto a nitrocellulose membrane (Hybond® ECL™, Merck, Germany)

using the WB technique. Transfer took place in transfer buffer at 1X dissolved in ddH₂O with methanol at 20 % (v/v). The voltage applied was 100 V for 2 hours on ice. Ponceau S (Sigma-Aldrich, Merck, Germany) staining was performed to confirm a successful transfer. Membrane blocking to prevent nonspecific binding of the antibodies was done in blocking buffer for 30 minutes on a shaker at room temperature. Immunodetection of proteins was done using primary antibodies (Table 3.2) at 1:1000 dilution in blocking buffer, overnight at 4 °C followed by 3 consecutive washes in TBS Tween 1X, each lasting 10 minutes. Incubation with secondary antibody (Table 3.2) at 1:5000 dilution in blocking buffer followed, for 1 hour on a shaker at room temperature. Another 3 washes in TBS Tween 1X, each lasting 10 minutes were done and the membrane was ready for protein detection. The enhanced chemiluminescence (ECL) Prime Western Blotting Detection Reagent (cat. number GERPN2232, GE Healthcare Life Sciences, USA) was used as a substrate for detection of proteins on the membrane. ImageQuant LAS 3000 (Fujifilm, Japan) system was used for capturing and digitalising images of bands on the membrane. ImageJ 1.x developed by Wayne Rasband in 1997 was used for band intensity quantification (Schneider *et al.*, 2012). All samples were normalised using WB results for housekeeping gene encoding actin.

3.4.10.4. Western blot membrane stripping

The process of stripping the nitrocellulose membrane is based on removal of the primary and secondary antibodies attached to it for further immunodetection of other proteins on the same blot. Two consecutive incubations in mild stripping buffer for 10 minutes on a shaker were performed after which the membrane was washed 3 times at room temperature with constant shaking in TBS Tween 1X, each wash lasting 10 minutes. Membrane blocking, followed by further incubations with primary and secondary antibodies and protein detection, as described above, was performed using different antibodies to detect other protein on the same membrane.

3.4.11. Extraction and purification of protein for activity assay

3.4.11.1. Extraction of glutathione S-transferase (GST)-tagged protein

Expression vector pGEX-6P-1 contains a sequence for a GST affinity tag and allows for production of a recombinant protein with the tag fused at its N-terminus (Appendix II). GST is a 211 amino acid (26 kDa) protein which allows for purification of proteins to which it is attached.

The wild-type and its variants were subcloned from plasmid *ADPRH* (BC063883.1) in pGEM-T into pGEX-6P-1 expression vector (Tables 3.1 and 3.3). In order to help counter protein misfolding and promote bacterial survival I used a plasmid expressing chaperonin GroEL/GroES (Goltermann *et al.*, 2016). These constructs were used for *E. coli DH5 α* co-transformation. After selection of the cells containing both plasmids by growth on a medium containing ampicillin (1:1000 of 100 mg/ml; AM04680005, Scharlab, Spain) and chloramphenicol (1:1000 of 33 mg/ml; C0378, Sigma-Aldrich, Merck, Germany), glycerol stocks were prepared, as described in section 3.3.4.

For extraction of GST-tagged protein the transformed *E. coli DH5 α* were grown in 3 ml of LB with ampicillin (1:1000 of 100 mg/ml; AM04680005, Scharlab, Spain) and chloramphenicol (1:1000 of 33 mg/ml; C0378, Sigma-Aldrich, Merck, Germany) overnight at 37 °C in a shaking incubator. The next day, 500 μ l of the cell culture were used for inoculation of 100 ml of LB with the corresponding antibiotics. The cells were grown overnight at 37 °C in a shaking incubator. The next day, 100 ml of the cell culture were used for inoculation of 500 ml of LB with the corresponding antibiotics. The cells were grown at 37 °C in a shaking incubator until optical density (OD) measured at 600 nm was between 0.6 and 0.8. Next, isopropyl β -D-1-thiogalactopyranoside (IPTG; cat. number GEN-S-02122, Generon, UK) was added to a final concentration of 0.8 mM for protein expression induction. The cell culture was left to grow at 37 °C in a shaking incubator for 4 hours. After incubation the cells were transferred onto ice and distributed into 50 ml falcons. Centrifugation at 2,600 x g for 10 minutes at 4 °C followed. After discarding the supernatant, the pellet was resuspended in 2 ml PBS 1X with a Protease Inhibitor Cocktail (cOmplete™, EDTA-free, Roche, Switzerland) 1X. The cells were collected in one 5 ml microcentrifuge tube and sonicated using Bioruptor (Diagenode, Belgium) on ice with 20 pulses of 20 seconds and 45 seconds pause in-between. After sonication the samples were centrifuged at 14,000 x g for 30 minutes at 4 °C. Both the supernatant and the pellet were stored at -20 °C until used.

3.4.11.2. Purification of GST-tagged protein

GSTrap FF 5 ml column (GE Healthcare Life Sciences, USA) for purification of GST-tagged proteins is a prepacked Glutathione Sepharose Fast Flow column. Protein antigenicity and function are preserved during the purification and the GST tag can be cleaved using PreScission protease (donation from Jeronimo Bravo laboratory). The

purification procedure using the GSTrap FF 5 ml column was done according to the providers' protocol, as follows:

1. Equilibration of column with 25 ml of PBS 1X.
2. Application of the supernatant from the extraction of GST-tagged protein using a syringe by pumping it onto the column.
3. Washing with 50 ml of PBS 1X.
4. Washing with 50 ml of PCB at pH 7.5.
5. Loading of the PreScission protease mixture (400 μ l (800 units) of PreScission protease with 4.6 ml of PCB at 4 °C) onto the column.
6. Incubation of the column at 4 °C for 4 hours.
7. Elution of the protein of interest with 15 ml of PCB leaving GST and PreScission protease bound to the column.

3.4.11.3. Sample dialysis

In order to substitute the PCB buffer with PBS 1X, the 15 ml of protein of interest from GSTrap FF 5 ml column purification was used in dialysis. Slide-A-Lyzer™ Dialysis Cassette (Thermo Fisher Scientific, USA) was used and the 15 ml of protein of interest were loaded onto the cassette with a syringe. The loaded dialysis cassette was dipped in PBS 1X for 2 hours at 4 °C on a magnetic stirrer. The PBS 1X was exchanged and the same incubation took place for another 2 hours. After that, the last PBS 1X exchange took place and the final incubation took place under the same conditions, this time overnight. The sample was recovered from the dialysis cassette the next day.

3.4.11.4. Protein concentration

Higher concentration of the protein of interest after dialysis was achieved using the Amicon® Ultra-15 Centrifugal Filter Units (Merck, Germany). The sample was loaded onto the filter and centrifuged at 4,000 x g for 20 minutes. Typically, 200 μ l of concentrated protein sample were recovered. The sample was stored at -20 °C until used.

3.4.12. Activity assay for ADPRH with cholera toxin

In order to assess the activity of ADPRH and variants of the protein, an assay was developed using Cholera Toxin B subunit (CT; Merck, Germany) as a catalyst of the ADP-ribosylation reaction. The assay was modified from a previously established method for

ADP-ribosyltransferase activity of CT (Suryadi and Shine, 2011). Agmatine sulfate salt (Sigma-Aldrich, Merck, Germany) was the ADP-ribose acceptor and β -nicotinamide adenine dinucleotide (β -NAD⁺) sodium salt (Merck, Germany) was used as the ADP-ribose donor in this arginine-specific reaction. The reaction of ADP-ribosylation took place in sodium phosphate with DTT as a reducing agent. The conditions of the reaction were as follows:

Agmatine sulfate salt	30 mM
β -NAD ⁺	7 mM
DTT	20 mM
pH 7.0 sodium phosphate	30 mM
CT	15 μ g/ml

The mixture was incubated at 37 °C for 90 minutes. Next, CT was removed from the mix using the Microcon-30 kDa Centrifugal Filter Unit with Ultracel-30 membrane (Merck, Germany). After ADP-ribosylation, ADP-ribose cleavage was initiated by addition of the ADPRH protein or its variants. These were obtained by transformation of *E. coli DH5 α* with *ADPRH* (BC063883.1), *ADPRH* p.R295P and *ADPRH* p.D55A/D56 all in pGEX-6P-1 vector with tac promoter for bacterial expression and purification as explained in section 3.4.11. 10 mM MgCl₂ and 15 μ g/ml of wild-type ADPRH protein or its variants were added to the mixture without CT and incubated for 90 minutes at 37 °C. The products of this reaction were loaded onto the ultraperformance liquid chromatography (UPLC) column. 5 μ l of each sample were injected into a UPLC-quadrupole time-of-flight mass spectrometry (UPLC-qTOF/MS) system equipped with an electrospray ionization (ESI) interface (LC, Acquity UPLC system, Waters, USA; MS, MICROMASS Q-ToF micro, Waters, USA) and a photodiode array detector. LC was carried out by reverse-phase ultraperformance liquid chromatography on a Waters BEH C18 UPLC column (1.7 micrometers) with dimensions: 2.1 x 150 mm. During sample running, the mobile phase consisted 0.1 % formic acid in water (phase A), and 0.1 % formic acid in acetonitrile (phase B). The solvent gradient program was 100 – 97.5 % solvent A over 8 min.

Passing the mixture through the column allows for separation, identification, and quantification of components based on data from a detector measuring absorbance at 260 nm. The UPLC procedure was performed by Ana Espinosa and Teresa Caballero Vizcaíno

from the Metabolomics Service at the Institute for Plant Molecular and Cell Biology (IBMCP-CSIC).

After establishment of a nicotinamide (Sigma-Aldrich, Merck, Germany) and β -NAD⁺ (Sigma-Aldrich, Merck, Germany) standard curves, the following reaction products were tested by loading onto UPLC:

1. Control reaction with all the substrates but without CT, ADPRH^{WT} nor its variants.
2. Reaction of the substrates with CT but without ADPRH^{WT} nor its variants.
3. Reaction of the substrates with CT followed by a reaction with ADPRH^{WT}.
4. Reaction of the substrates with CT followed by a reaction with ADPRH^{R295P}.
5. Reaction of the substrates with CT followed by a reaction with ADPRH^{D55A/D56A}.

3.4.13. Co-immunoprecipitation (co-IP)

Protein-protein interactions can be identified by co-IP. Indirect capture of proteins bound to the target protein is possible through target protein specific antibodies. In this study the first step to verify an interaction between ADPRH and ADPRHL1 was to co-transfect SH-SY5Y cells with both *ADPRH* (BC063883.1) in pFlag and *ADPRHL1* (NM_138430.5) in pcDNA3.1+/C-(K)DYK. Separate transfections of each construct and a mock transfection were performed for control. All transfections were performed as described in section 3.4.7.2. After standard incubation of 48 hours the proteins were extracted as described in section 3.4.10.1. Co-IP protocol followed, using Protein G Mag Sepharose Xtra (GE Healthcare Life Sciences, USA). A 100 μ l of magnetic beads slurry was pipetted into a 1.5 ml tube and placed in a magnetic rack. The medium was discarded and the beads were resuspended in 500 μ l PBS 1X. The slurry was placed in the magnetic rack once again, the PBS 1X was discarded and then 4 μ l of antibody in 296 μ l PBS 1X was added. The antibodies used were: Anti-ADPRH antibody produced in rabbit and Monoclonal ANTI-FLAG® M1 antibody produced in mouse. The tubes were left overnight at 4 °C on a shaker. The next day the antibody in PBS 1X was removed on the magnetic rack and 300 μ l of the protein extract was added. The tubes were left overnight at 4 °C on a shaker. The following day the protein extract was removed and three washes using PBS 1X were performed consisting of resuspending the beads in PBS and its removal with the tubes on the magnetic rack. Elution was performed using 100 μ l of glycine 100 mM (PanReac AppliChem, USA) adjusted to pH 2.8 using HCl (Scharlab, Spain). Elution took 10 minutes.

The resulting protein solution was quantified and used for WB as described in sections: 3.4.10.2. and 3.4.10.3. To confirm the target protein has been obtained in the eluate the membranes were then stripped and another WB was performed using appropriate antibodies (section 3.4.10.4.).

3.4.14. Statistical analysis

Statistical analysis was performed using the student's t test and one-way or two-way ANOVA to test differences between group means, with a *post hoc* Tukey multiple comparisons of means test. Two-way ANOVA was used when data consisted of two independent variables. In all tests the level of significance was set at $p < 0.05$. All data are presented as mean \pm standard error of the mean (SEM). All statistical analyses were performed using the 3.6.1 version of R software developed by John Chambers and colleagues at Bell Laboratories (R Core Team, 2019).

4. RESULTS

4.1. Study of etiopathogenesis of AD in family UGM037

4.1.1. Clinical data

Three members of family UGM037 were diagnosed with AD by computed tomography scan. Typical clinical features of AD were observed in the patients. Initial symptoms were memory loss and a progressive cognitive decline. The clinical data on the three affected individuals and their sibling are summarised in table 4.1.

Table 4.1. Data including diagnosis, age of onset and age of death for family UGM037 affected members and their sibling.

Sample ID	Sex	Diagnosis	Age of onset	Age and cause of death
04-046	Female	AD	69 y.o.	88 y.o., AD related
04-047	Male	non-AD		88 y.o., Cancer, no dementia detected
04-048	Male	AD	72 y.o.	85 y.o., AD related
04-049	Female	AD	68 y.o.	89 y.o., AD related

The patients all had onset of AD at an age above 68 years. The children of patients and the unaffected member of the family, whose DNA samples have been collected at ages between 50 and 57 (Figure 3.1), did not show any symptoms.

4.1.2. WES data

4.1.2.1. Base-calling and image analysis raw data

Basic raw data obtained through base-calling and image analysis produced by WES for samples 04-046 and 04-048 are summarised in table 4.2.

Table 4.2. WES raw data: reads, base content and phred quality scores for samples 04-046 and 04-048.

Sample ID	Total read bases (bp)	Total reads	GC (%)	Q20 (%)	Q30 (%)
04-046	8,610,828,932	85,255,732	49.7	98.8	96.5
04-048	9,659,430,728	95,637,928	49.5	98.8	96.5

Table 4.2. WES raw data: reads, base content and phred quality scores for samples 04-046 and 04-048 (continued).

Total read bases – Total number of bases sequenced; Total reads – Total number of reads. In illumina paired-end sequencing, read1 and read2 are added; GC (%) – GC content; Q20 (%) – Ratio of reads that have phred quality score of over 20; Q30 (%) – Ratio of reads that have phred quality score of over 30.

4.1.2.2. Read alignment results

The average read length for both samples was 101 bp. The post-alignment statistics are given in table 4.3.

Table 4.3. WES post-alignment statistics for samples 04-046 and 04-048.

Sample ID	04-046	04-048
Initial Mappable Reads	85,163,923	95,499,933
% Initial Mappable Reads	99.8	99.8
Non-Redundant Reads	75,137,432	83,306,002
% Non-Redundant Reads	88.2	87.2
On-Target Reads	62,211,853	68,805,631
% On-Target Reads	82.7	82.5
On-Target Yield (bp)	5,450,617,779	6,012,147,064
Mean Depth of Target Regions (X)	108.1	119.3

Initial Mappable Reads – Number of mapped reads to human genome; % Initial Mappable Reads – $100 * (\text{Initial mappable reads}) / (\text{Total reads})$; Non-Redundant Reads – Number of de-duplicate reads from Picard tools; % Non-Redundant Reads – $100 * (\text{Non-redundant reads}) / (\text{Initial mappable reads})$; On-Target Reads – Number of reads mapped to target regions; % On-Target Reads – $100 * (\text{On-target reads}) / (\text{Non-redundant reads})$; On-Target Yield (bp) – The sum of the bases in the final alignment to the target regions; Mean Depth of Target Regions (X) – $(\text{On-target yield}) / (\text{Target regions})$

4.1.2.3. SNP calling results

SNP calling allowed for identification of variants in DNA sample sequences. The SNP calling results for samples 04-046 and 04-048 are summarised in table 4.4.

Table 4.4. WES SNP calling summary for samples 04-046 and 04-048.

Sample ID	04-046	04-048
Number of SNPs	78,134	79,669
Synonymous Variants	11,396	11,404
Missense Variants	10,336	10,363
Stop Gained	103	93
Stop Lost	38	41
Number of INDELS	9,120	9,583
Frameshift Variants	279	279
Inframe Insertions	158	150
Inframe Deletions	175	177
% Found in dbSNP142	97.5	97.3
Het/Hom Ratio	1.5	1.5
Ts/Tv Ratio	2.3	2.3

% Found in dbSNP142 – percentage of Number of SNPs found in dbSNP142; Het/Hom Ratio – Ratio of number of heterozygous variants to number of homozygous variants; Ts/Tv Ratio – Ratio of transition rate of SNVs that pass the quality filters divided by transversion rate of SNVs that pass the quality filters. Transitions are interchanges of purines (A, G) or of pyrimidines (C, T). Transversions are interchanges between purine and pyrimidine bases.

4.1.3. WES results filtration and identification of interesting variants

The results obtained from WES analysis needed to be filtered and the most interesting variants assessed. Firstly, genes known to be associated with AD were looked at, to see whether a known or a new variant in those genes could cause or affect the disease. The summary of findings can be found in table 4.5.

Table 4.5. WES results filtered for AD associated genes for samples 04-046 and 04-048.

Gene	Number of intron variants	Number of synonymous variants	Number of missense variants	Other variants
<i>ABCA7</i>	13	6	3	0
<i>ADAM10</i>	5	0	0	0
<i>AKAP9</i>	14	6	3	inframe insertion
<i>APOE</i>	1	0	1	0
<i>APP</i>	9	0	0	0
<i>BIN1</i>	1	0	0	0
<i>CASS4</i>	2	5	1	3'-UTR variant
<i>CD2AP</i>	3	1	0	5'-UTR variant
<i>CELF1</i>	1	0	0	0
<i>CLU</i>	3	1	0	0
<i>CR1</i>	2	1	3	stop gain variant
<i>DSG2</i>	3	3	0	0
<i>EPHA1</i>		2	2	0
<i>FERMT2</i>	2	1	0	3'-UTR variant
<i>INPP5D</i>	17	2	1	5'-UTR variant
<i>NME8</i>	7	3	1	0
<i>PICALM</i>	6	2	0	0
<i>PLD3</i>	2	0	0	3'-UTR variant
<i>PSEN1</i>	1	0	0	0
<i>PSEN2</i>	8	3	0	0
<i>PTK2B</i>	22	4	0	0
<i>SORL1</i>	13	5	2	0
<i>UNC5C</i>	5	4	1	0
<i>ZCWPW1</i>	2	0	0	0

Of the variants in table 4.5, only one had a putative impact predicted as HIGH by SnpEff software. It was the variant rs764542666 in gene *CRI*. This variant encodes a PTC c.C406T p.R136*. Other variants in table 4.5 were filtered out from further analysis either by prediction of their impact, their frequency in population or a review of current knowledge on their effect.

In parallel, all variants, whether related to AD or not, went through a restrictive filtering process to assess the possible cause of AD in family UGM037. Filtering was done as described in the methods section. An important step was the filtering by role, function, tissue of expression and pathways the variants are involved in. The other variants found through filtering are summarised in table 4.6.

This separate filtering, which considered involvement in AD related pathways to eliminate variants at the last stage, once again suggests the rs764542666 in gene *CRI* may be important in causing the disease. Out of the seven variants in table 4.6, only rs764542666 exists in the dbSNP142 and other databases used. All other are novel variants, which is partly why they were included in this list.

The novel variants in genes *ANKRD36*, *GXYLT1* and *CERS5* encoded PTCs. *ANKRD36* is associated with giant axonal neuropathy (Johnson-Kerner *et al.*, 2015), which is a disease affecting the central and peripheral nervous systems, most typically in children below five years of age. *GXYLT1* is associated with skin diseases and *CERS5* with endocrine diseases (Stelzer *et al.*, 2016; www.genecards.org). After further literature review, these three variants could also be discarded.

KRT9 is considered a biomarker for AD (Richens *et al.*, 2016). While it is suggested *KRT9* expression is dysregulated as a consequence of AD pathology, the relation has not been fully studied thus making this an interesting variant to study. A known variant at the same amino acid encoding the following change p.D238Y is predicted to be deleterious by SIFT and probably damaging by Polyphen (Karczewski *et al.*, 2020; <https://gnomad.broadinstitute.org/>).

ADAM family of proteins have been associated with AD (Moss *et al.*, 2001), however, ADAM33 has been linked with asthma and not AD, thusfar (Van Eerdewegh *et al.*, 2002). Polymorphisms in *PRND* have a more direct correlation with AD (Schröder *et al.*, 2001; Flirski *et al.*, 2012).

Table 4.6. WES results filtered for possibly damaging variants for samples 04-046 and 04-048.

Chr	Pos	Ref	Alt	Zigosity	Gene	Amino acid variant	SIFT prediction (score)	Polyphen prediction (score)	Putative Impact (SnPEff)
chr1	207,684,918	C	T	HET	<i>CR1</i>	p.R136*	-	-	HIGH
chr2	97,851,077	C	G	HET	<i>ANKRD36</i>	p.S655*	-	-	HIGH
chr12	42,538,352	C	A	HET	<i>GXYLT1</i>	p.G33*	-	-	HIGH
chr12	50,528,463	C	A	HET	<i>CERS5</i>	p.E299*	-	-	HIGH
chr17	39,726,402	T	G	HET	<i>KRT9</i>	p.D238A	D (0.001)	D (1.0)	MODERATE
chr20	3,654,252	C	G	HET	<i>ADAM33</i>	p.C320S	D (0.003)	D (1.0)	MODERATE
chr20	4,705,588	G	A	HET	<i>PRND</i>	p.V131M	D (0.002)	D (0.986)	MODERATE

Chr – chromosome; Pos – position; Ref – reference allele; Alt – alternate allele; HET – heterozygous; D – deleterious (SIFT)/damaging (Polyphen). SIFT and Polyphen both do not perform predictions for nonsense variants.

The standards and guidelines for the interpretation of sequence variants (Richards *et al.*, 2015) and the gene constraint observed/expected score (Karczewski *et al.*, 2020; <https://gnomad.broadinstitute.org/>) were used to assess these variants. Low values of the observed/expected score are indicative of strong intolerance. A summary can be seen in table 4.7.

Table 4.7. Interpretation of the selected variants according to Richards *et al.* (2015) and observed/expected score.

Gene	Amino acid variant	Effect	Sequence variant classification	observed/expected score (90 % CI)
<i>CRI</i>	p.R136*	stop gained	Likely pathogenic	0.78 (0.64 - 0.96)
<i>KRT9</i>	p.D238A	missense variant	Uncertain significance	0.97 (0.89 - 1.06)
<i>ADAM33</i>	p.C320S	missense variant	Uncertain significance	0.91 (0.84 - 0.99)
<i>PRND</i>	p.V131M	missense variant	Uncertain significance	1.03 (0.88 - 1.21)

4.1.4. Sanger sequencing verification of selected variants

The variants selected from WES results needed to be confirmed by Sanger sequencing to eliminate a possibility of a false positive result. As only samples 04-046 and 04-048 were used in WES, samples 04-047 and 04-049 were added to the Sanger sequencing confirmation. Since 04-049 was known to come from an AD patient and 04-047 from a non-AD member of the family, this would also be a way to see whether a variant segregates with the disease (Table 4.8).

Table 4.8. Verification of selected variants through Sanger sequencing in family UGM037.

Gene	Amino acid variant	Ref	Alt	Sample ID and diagnosis			
				04-046	04-047	04-048	04-049
<i>CR1</i>	p.R136*	C	T	C/T	C/C	C/T	C/T
<i>KRT9</i>	p.D238A	T	G	T/G	T/T	T/G	T/G
<i>ADAM33</i>	p.C320S	C	G	C/G	C/G	C/G	C/G
<i>PRND</i>	p.V131M	G	A	G/A	G/A	G/A	G/G

Ref – reference allele; Alt – alternate allele

Sanger sequencing confirms the WES results, and shows segregation of variants in genes *CR1* and *KRT9* with the disease. Further Sanger sequencing of samples 19-059 – 19-065, taken from children of the subjects was done to establish the frequency of the variant rs764542666 within the family. Even though these family members did not show AD symptoms, they are at a risk of developing the disease at an older age. Figure 4.1 shows the UGM037 family tree with the genotype of each member, whose DNA sample was available, for rs764542666 variant.

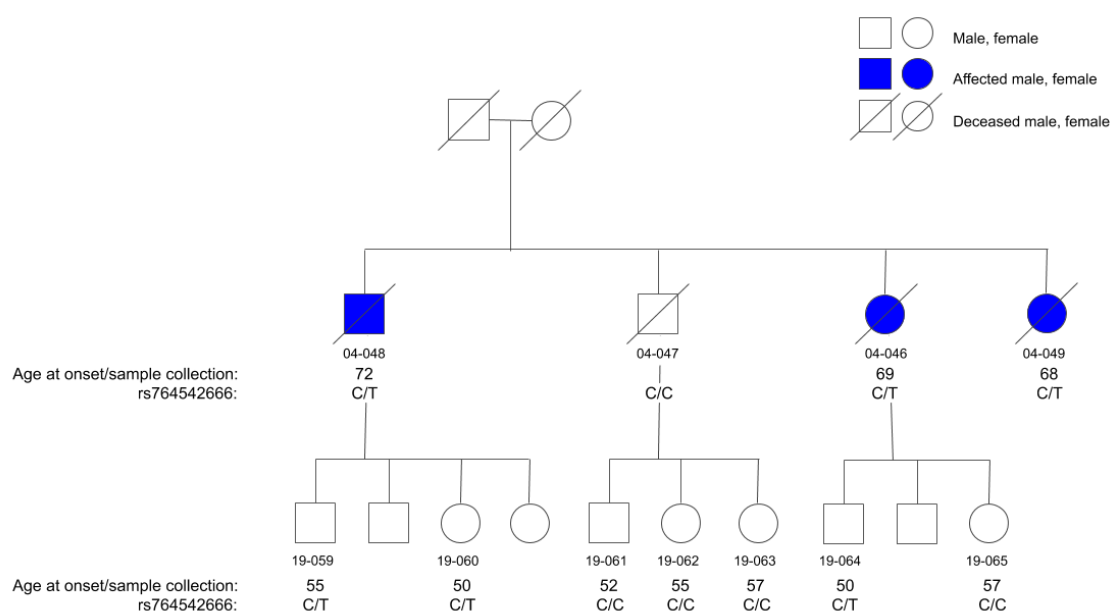


Figure 4.1. UGM037 family tree with the three AD affected members denoted by coloured in shapes. Genotypes for rs764542666 variant are included.

4.1.5. Frequency of the rs764542666 variant in control population

Frequency of a variant in a general population is an important factor in determining its pathogenicity. According to gnomAD (Karczewski *et al.*, 2020; <https://gnomad.broadinstitute.org/>) the frequency of rs764542666 is 0.00002008 (Table 4.9).

Table 4.9. General population frequency for rs764542666 with data on specific populations.

Population	Allele count	Allele number	Number of homozygotes	Allele frequency
Latino	3	34464	0	8.705×10^{-5}
European (non-Finnish)	2	112894	0	1.772×10^{-5}
African	0	15482	0	0.000
Ashkenazi Jewish	0	10060	0	0.000
East Asian	0	17974	0	0.000
European (Finnish)	0	21520	0	0.000
Other	0	6046	0	0.000
South Asian	0	30600	0	0.000
Female	3	113946	0	2.633×10^{-5}
Male	2	135094	0	1.480×10^{-5}
Total	5	249040	0	2.008×10^{-5}

(data from Karczewski *et al.*, 2020; <https://gnomad.broadinstitute.org/>)

In addition to obtaining the frequency of the variant from public databases of genetic variations, ASPCR was used to scan a control population available in the laboratory for rs764542666. Additionally, a population of AD patients was scanned for this variant. The data is summarised in table 4.10. Members of the family who may be presymptomatic are not included.

Table 4.10. Number of cases of rs764542666 variant in different groups.

Group	CR1 p.R136* C>T		
	CC	CT	TT
Spanish control population	192	0	0
Alzheimer's disease patients	124	0	0
Unaffected family members	1	0	0
Affected family members	0	3	0

4.1.6. Genotyping *APOE* and *CR1* polymorphisms

The $\epsilon 4$ allele of *APOE* is a well-established risk factor for AD. Genotyping this gene in UGM037 family was important to establish whether it possibly contributed to the familial disease. The number of $\epsilon 4$ alleles present are directly related to earlier onset and higher severity of the disease (Giri *et al.*, 2016). *CR1* is also a known genetic risk factor for AD. To explore the possibility of the rs764542666 variant being in linkage disequilibrium with a known risk allele, we genotyped the rs3818361 SNP and the long isoform of *CR1* (*CR1**2) polymorphisms, which were both described as contributing to AD susceptibility (Mahmoudi *et al.*, 2015; Li *et al.*, 2016). In rs3818361, an adenine instead of a guanine is considered a risk factor, while the *CR1**2 isoform has been associated through GWAS with AD. The genotyping was done to reveal whether the genotypes relate to diagnosis. The results are presented in figure 4.2.

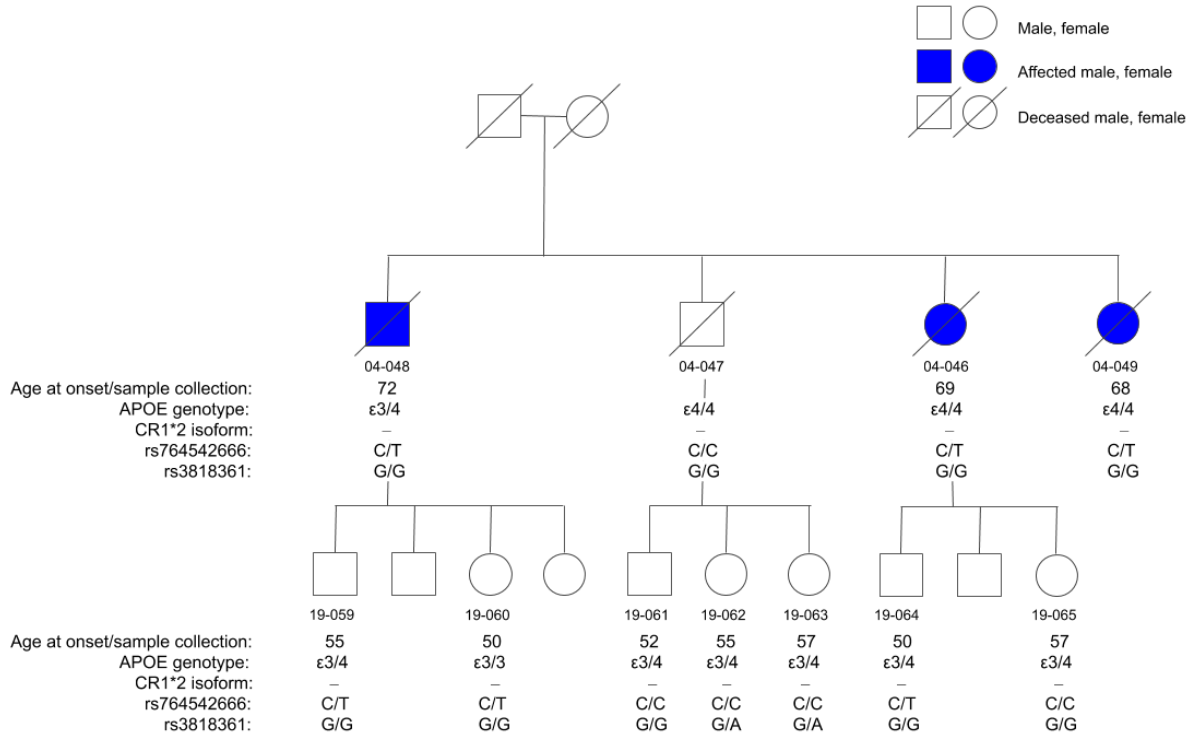


Figure 4.2. UGM037 family tree with the three AD affected members denoted by coloured in shapes. Genotypes for rs764542666 and rs3818361 variants as well as APOE and CR1*2 isoform are included.

“-” – Lack of isoform CR1*2.

4.1.7. Role of variant rs764542666 (p.R136*) in *CR1* mRNA expression

To date, there has been no mention of a mutation in *CR1* causing familial AD. Variants in *CR1* have only been associated with a higher risk of developing the disease. In particular, no known variant encoding a PTC in *CR1* has been found to cause AD. Wild-type *CR1* mRNA was used for comparison. Transfection efficiency was controlled by amplification of kanamycin from the plasmid resistance cassette.

The variant CR1^{R136*} introduced a stop codon early on in the 2039 amino acid protein. Analysis of mRNA levels using the plasmids carrying the CR1^{WT} gene and the CR1^{R136*} variant was done using RT-qPCR method and SH-SY5Y cells. The results show a significant 5.4-fold decrease ($p = 0.0004$) in CR1^{R136*} variant's mRNA levels in comparison with the wild-type (Figure 4.3).

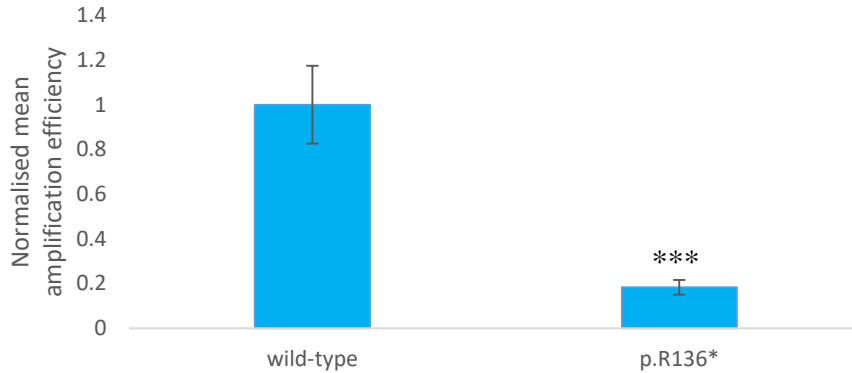


Figure 4.3. RT-qPCR results for mRNA levels of CR1^{WT} and CR1^{R136*} variant, overexpressed in SH-SY5Y cells.

*** $p < 0.001$ student's t test comparing CR1^{WT} and its variant. Data obtained from three technical and two biological replicates.

The mean amplification efficiency was normalised to a hundred percent for the CR1^{WT} (Mean \pm SEM: 1.000 \pm 0.174). The result for CR1^{R136*} variant was 5.4 times lower (0.184 \pm 0.033). NMD pathway was inhibited by NMDI14 in an experimental group to verify whether this would impact the mRNA levels of CR1^{R136*}. After NMDI14 treatment the mean mRNA levels of CR1^{R136*} increased 3.4 fold with respect to untreated CR1^{R136*} (Figure 4.4).

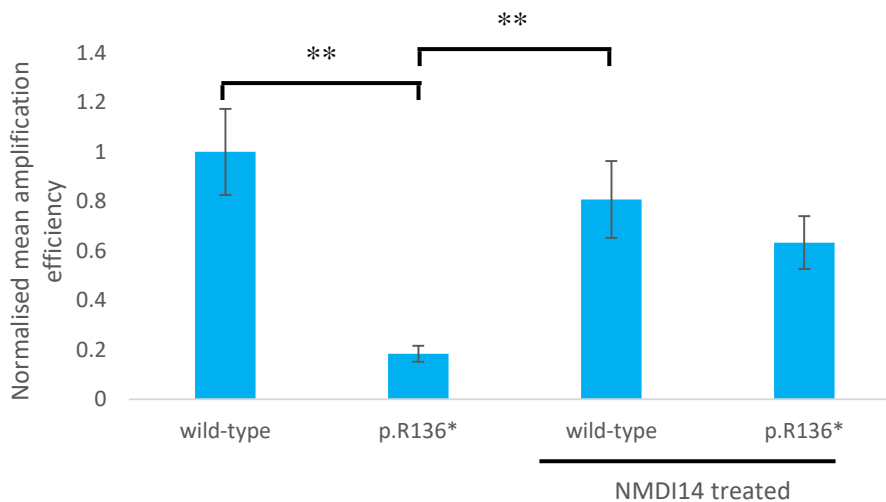


Figure 4.4. RT-qPCR results for mRNA levels of CR1^{WT} and CR1^{R136*} variant overexpressed in SH-SY5Y cells with and without NMD pathway inhibition.

** $p < 0.01$ two-way ANOVA comparing CR1^{WT} and its variant, with *post hoc* Tukey multiple comparisons of means test. Data obtained from three technical and two biological replicates.

The mRNA levels of CR1^{R136*} variant increased from 0.184 ± 0.033 to 0.633 ± 0.107 (Mean \pm SEM). The statistically significant difference between CR1^{WT} and CR1^{R136*} in the non-NMDI14 treated group disappear in the NMDI14 treated group where the mRNA levels of both are similar ($p = 0.775$).

Comparison of the increase of mRNA levels between the variant and the wild-type after NMDI14 treatment showed a significant 3.4-fold increase for CR1^{R136*} ($p = 0.01$; Figure 4.5). The CR1^{WT} mRNA levels decreased by approximately 0.8 fold.

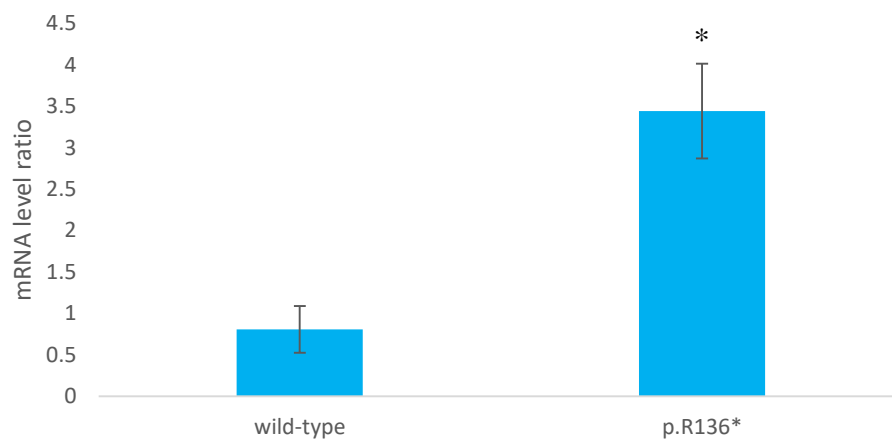


Figure 4.5. Ratio of mRNA levels change after treatment with NMDI14 for CR1^{WT} and CR1^{R136*} variant overexpressed in SH-SY5Y cells.

* $p < 0.05$ student's t test comparing CR1^{WT} and its variant. Data obtained from three technical and two biological replicates.

4.1.8. Role of variant rs764542666 (p.R136*) in CR1 protein expression

Analysis of protein overexpression using the plasmids carrying the CR1^{WT} gene and the CR1^{R136*} variant was analysed by WB in SH-SY5Y cells. The results show a significant ($p = 0.0097$) decrease in CR1^{R136*} variant protein levels in comparison with the wild-type (Figure 4.6).

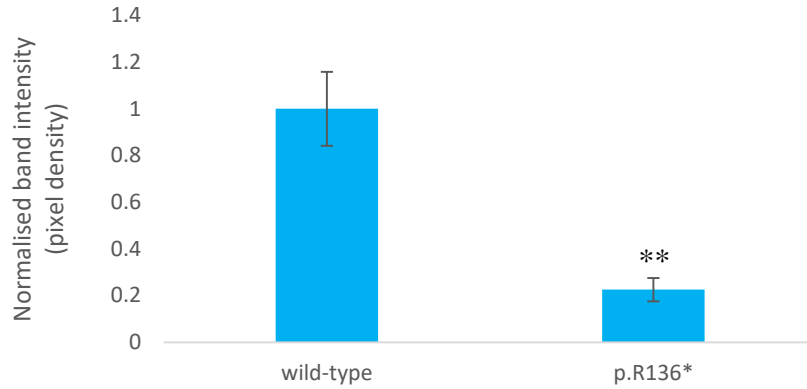


Figure 4.6. WB results for protein levels of CR1^{WT} and CR1^{R136*} variant, overexpressed in SH-SY5Y cells.

** $p < 0.01$ student's t test comparing CR1^{WT} and its variant. Data obtained from three technical and biological replicates.

The band intensity was normalised for the CR1^{WT} (Mean \pm SEM: 1.000 \pm 0.159). The result for CR1^{R136*} variant was 4.4 times lower (0.226 \pm 0.050). To study whether NMD pathway inhibition would increase the protein levels of CR1^{R136*}, NMDI14 was used. A significant increase in CR1^{R136*} protein levels from 0.226 \pm 0.050 to 0.778 \pm 0.118 was observed ($p < 0.05$; Figure 4.7).

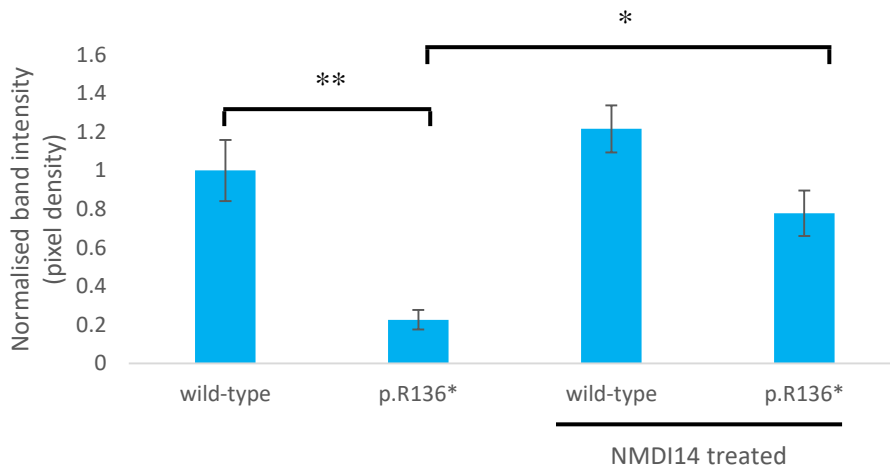


Figure 4.7. WB results for protein levels of CR1^{WT} and CR1^{R136*} variant overexpressed in SH-SY5Y cells with and without NMD pathway inhibition.

* $p < 0.05$, ** $p < 0.01$ two-way ANOVA comparing CR1^{WT} and its variant, with *post hoc* Tukey multiple comparisons of means test. Data obtained from three technical and biological replicates.

The increase in CR1^{WT} protein levels from 1 ± 0.159 to 1.216 ± 0.122 after NMDI14 treatment was not significant. In summary, CR1^{R136*} protein levels were lower than those of CR1^{WT} (treated or untreated with NMDI14) and increased significantly after inhibition of the NMD pathway.

Comparison of CR1^{WT} and CR1^{R136*} increase of protein levels after NMD pathway inhibition, proved to be significant ($p = 0.013$). The protein levels for CR1^{R136*} increased 3.6 fold and only 1.3 fold for CR1^{WT} (Figure 4.8).

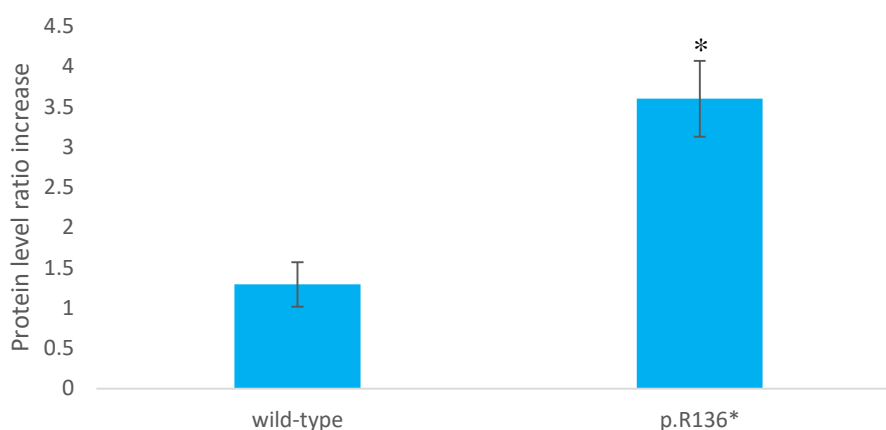


Figure 4.8. Ratio of protein levels increase after treatment with NMDI14 for CR1^{WT} and CR1^{R136*} variant overexpressed in SH-SY5Y cells.

* $p < 0.05$ student's t test comparing CR1^{WT} and its variant. Data obtained from three technical and biological replicates.

4.2. Study of etiopathogenesis of PLS in family UGM471

4.2.1. Clinical data

Three patients were diagnosed with PLS in family UGM471. A female and two of her sons. At 30 years of age the mother (for privacy purposes assigned number 0002 in figure 3.2) was showing a feeling of stiffness in one leg which later spread to the second leg and upper extremities. Further, her facial muscles were affected, her voice was altered and problems with deglutition appeared. Forgetfulness was followed by a complete loss of speech. The progression of the disease lasted 16 years. The most recent report described her as bedridden, with rigidity and unable to move her limbs. Minor eye movement can be observed. Fed through percutaneous endoscopic gastrostomy (PEG) tube, but mechanical ventilation is not necessary.

Patients 1001 and 1002, the sons, were also diagnosed with PLS. The first, was diagnosed in 2005 at 24 years of age while the second was diagnosed in 2013 at 28 years of age. Initial symptom of a feeling of stiffness in left leg was observed in the younger patient 1002. The clinical data on the three affected individuals are summarised in table 4.11.

Table 4.11. Data including diagnosis and age of onset for family UGM471 affected members.

Sample ID	Sex	Diagnosis	Age of onset
0002	Female	PLS	30 y.o.
1001	Male	PLS	24 y.o.
1002	Male	PLS	28 y.o.

4.2.2. WES data

The DNA samples of the three patients were sequenced using HiSeq 2000[®] sequencer (Sistemas Genomicos, Valencia, Spain) unlike the other samples in this work which were sequenced using HiSeq 4000[®] sequencer (Macrogen, Seoul, Rep. of Korea). Any discrepancies between the results are due to this difference and differences in the way the companies provide result reports.

4.2.2.1. Base-calling and image analysis raw data

Basic raw data obtained through base-calling and image analysis produced by WES for samples 0002, 1001 and 1002 are summarised in table 4.12.

Table 4.12. WES raw data for samples 0002, 1001 and 1002.

Sample ID	Total reads	% mapped	% HQ mapped	% HQ mapped no PCR dups
0002	47,552,618	99.13	97.68	94.78
1001	56,331,030	98.75	97.89	92.96
1002	51,859,032	99.1	97.64	93.87

Total reads – total number of generated reads; % mapped – percentage of mapped reads; % HQ mapped – % reads after low quality removal; % HQ mapped no PCR dups – percentage of reads that remain after removing read PCR duplicates.

4.2.3. WES results filtration and identification of interesting variants

Again the results obtained from WES analysis needed to be filtered and the most interesting variants assessed. Firstly, genes known to be associated with PLS and ALS were looked at, to see whether a known or a new variant in those genes could cause or affect the disease. The summary of findings can be found in table 4.13.

Table 4.13. WES results filtered for PLS associated genes for samples 0002, 1001 and 1002.

Gene	Number of intron variants	Number of synonymous variants	Number of missense variants	Other variants
<i>ALS2</i>	4	1	1	0
<i>ATXN2</i>	-	2	1	0
<i>C9orf72</i>	1	0	0	3'-UTR variant
<i>CHCHD10</i>	1	2	0	0
<i>CHMP2B</i>	0	1	0	0
<i>ERBB4</i>	3	1	1	0
<i>FIG4</i>	4	1	1	3'-UTR variant splicing variant
<i>FUS</i>	1	1	0	5'-UTR variant
<i>GLE1</i>	2	0	1	3'-UTR variant
<i>NEFH</i>	1	3	0	nonframeshift insertion
<i>OPTN</i>	6	0	1	splicing variant
<i>SETX</i>	6	2	6	3'-UTR variant
<i>SIGMAR1</i>	0	0	1	3'-UTR variant
<i>SPG7</i>	5	0	0	0
<i>SPG11</i>	2	0	1	0
<i>SSI8L1</i>	1	1	0	0
<i>TBK1</i>	3	0	0	0
<i>VCP</i>	3	0	0	splicing variant

No common variants were found in genes *SOD1* nor *TARDBP* in all of the affected individuals. The variants in table 4.13 had high frequencies in the general population and SIFT and Polyphen predictions pointed to their benign or tolerant impact.

Other variants, not associated with ALS or PLS were filtered by a restrictive process to assess the possible cause of the disease in family UGM471. Filtering was done as described in the methods section. An important step was the filtering by role, function, tissue of expression and pathways the variants are involved in. The other variants found through filtering are summarised in table 4.14.

Table 4.14. WES results filtered for possibly damaging variants for samples 0002, 1001 and 1002.

Chr	Pos	Ref	Alt	Zigosity	Effect	Gene	Amino acid variant	SIFT prediction (score)	Polyphen prediction (score)
chr1	154,842,244	A	T	HET	missense variant	<i>KCNN3</i>	p.L66H	D (0.01)	D (1.0)
chr3	119,306,535	G	C	HET	missense variant	<i>ADPRH</i>	p.R295P	D (0.01)	D (0.999)
chr7	70,229,818	C	A	HET	missense variant	<i>AUTS2</i>	p.P432H	T (0.17)	D (1.0)
chr9	139,413,097	T	G	HET	missense variant	<i>NOTCH1</i>	p.T349P	D (0.01)	D (0.999)
chr9	140,777,306	C	G	HET	missense variant	<i>CACNA1B</i>	p.N167K	D (0.02)	D (1.0)
chr10	75,006,782	T	G	HET	missense variant	<i>DNAJC9</i>	p.T56P	D (0.0)	D (1.0)

Chr – chromosome; Pos – position; Ref – reference allele; Alt – alternate allele; HET – heterozygous; D – deleterious (SIFT)/damaging (Polyphen); T – tolerated

KCNN3 is Potassium Calcium-Activated Channel Subfamily N Member 3 which is a gene with homeostatic role. While no direct association with ALS or PLS has been described, riluzole, the first systemic therapy drug for ALS, was found to activate *KCNN3* channels (Grunnet *et al.*, 2001). As stated in the introduction, *ADPRH* has been implicated

in cell proliferation, tumorigenesis, intracellular signal transduction and cell cycle regulation (Kato *et al.*, 2011). *AUTS2* is associated with certain neurodevelopmental and neurological disorders (provided by RefSeq, O'Leary *et al.*, 2016). NOTCH1 is part of the NOTCH family of proteins involved in intracellular signalling which has been found to regulate neuronal survival (Gómez-Pinedo *et al.*, 2019). *CACNA1B* is involved in control of neurotransmitter release. DNAJC9 is part of the DNAJ family of molecular chaperones. Members of this family have been implicated in neurodegenerative diseases (Stelzer *et al.*, 2016; www.genecards.org).

No nonsense variants of interest were identified within the WES results. The standards and guidelines for the interpretation of sequence variants (Richards *et al.*, 2015) were difficult to apply here as none of these genes are associated with ALS or PLS and ADPRH^{R295P} and DNAJC9^{T56P} are two previously unknown variants. They could all be classified as uncertain significance. The gene constraint observed/expected score (Karczewski *et al.*, 2020; <https://gnomad.broadinstitute.org/>) was used to assess these variants, where low values are indicative of strong intolerance. The data is summarised in table 4.15.

Table 4.15. Interpretation of the selected variants according to the observed/expected score.

Gene	Amino acid variant	Effect	observed/expected score (90 % CI)
<i>ADPRH</i>	p.R295P	missense variant	0.96 (0.85 - 1.08)
<i>AUTS2</i>	p.P432H	missense variant	0.78 (0.73 - 0.83)
<i>CACNA1B</i>	p.N167K	missense variant	0.65 (0.61 - 0.69)
<i>DNAJC9</i>	p.T56P	missense variant	0.97 (0.84 - 1.12)
<i>KCNN3</i>	p.L66H	missense variant	0.55 (0.50 - 0.62)
<i>NOTCH1</i>	p.T349P	missense variant	0.76 (0.73 - 0.80)

4.2.4. Sanger sequencing verification of selected variants and result assessment

False positive WES results could be eliminated using Sanger sequencing. Also, the DNA samples of non-PLS members of family UGM471 were tested for further exclusion of variants which did not segregate with the disease (Table 4.16).

Table 4.16. Verification of selected variants through Sanger sequencing in family UGM471.

Gene	Amino acid variant	Ref	Alt	Sample ID and diagnosis						
				9996 non-PLS	0001 non-PLS	0002 PLS	0003 non-PLS	1001 PLS	1002 PLS	1003 non-PLS
<i>ADPRH</i>	p.R295P	G	C	G/G	G/G	G/C	G/C	G/C	G/C	G/C
<i>AUTS2</i>	p.P432H	C	A	C/C	C/C	C/A	C/C	C/A	C/A	C/C
<i>CACNA1B</i>	p.N167K	C	G	–	–	C/C	–	C/C	C/C	–
<i>DNAJC9</i>	p.T56P	T	G	T/T	T/T	T/T	T/T	T/T	T/T	T/T
<i>KCNN3</i>	p.L66H	A	T	A/T	A/A	A/T	A/T	A/T	A/T	A/T
<i>NOTCH1</i>	p.T349P	T	G	–	–	T/T	–	T/T	T/T	–

Ref – reference allele; Alt – alternate allele; “–” – not sequenced

Sanger sequencing confirmed WES results for genes *KCNN3*, *ADPRH* and *AUTS2*. However, the results for *DNAJC9*, *CACNA1B* and *NOTCH1* were false positive. Variants in genes *KCNN3*, *ADPRH* and *AUTS2* are all present in the patients, however, only *AUTS2*^{P432H} was not detected in non-PLS members of the family. The variant *AUTS2*^{P432H} can be found in the dbSNP142 database with the accession number rs767529359. The ClinVar (Landrum *et al.*, 2018; <https://www.ncbi.nlm.nih.gov/clinvar/>) interpretations of pathogenicity for this variant are conflicting with the latest assessment suggesting it is benign (accession RCV000626221.2). Also, as mentioned, mutations in *AUTS2* are related to neurodevelopmental and neurological disorders, such as autism spectrum disorders or mental retardation and not to neurodegenerative diseases.

Variant *KCNN3*^{L66H} is also present in the dbSNP142 database, with the accession number rs776143138. While it is not reported in ClinVar, it has an estimated frequency of 0.02432 (837/34422) according to gnomAD (Karczewski *et al.*, 2020; <https://gnomad.broadinstitute.org/>). It also appears to be present in most of the family members whose DNA samples were available.

Variant *ADPRH*^{R295P} has not been previously described and is not present in any variant database. However, variants in *ADPRH* at the same nucleotide or amino acid position have been described. They are summarised in table 4.17.

Table 4.17. Summary of known variants in *ADPRH* at amino acid position 295.

Amino acid variant	Nucleotide change	Reference SNP accession number	Population frequency (allele count/allele number)
p.R295Q	c.G884A	rs150984649	9.194×10^{-5} (26/282,798)
p.R295L	c.G884T	rs150984649	3.185×10^{-5} (1/31,398)
p.R295*	c.C883T	rs190409478	5.269×10^{-4} (149/282,764)
p.R295P	c.G884C	-	-

Figure 4.9 shows the UGM471 family tree with the genotype of each member whose DNA sample was available for *ADPRH*^{R295P} variant.

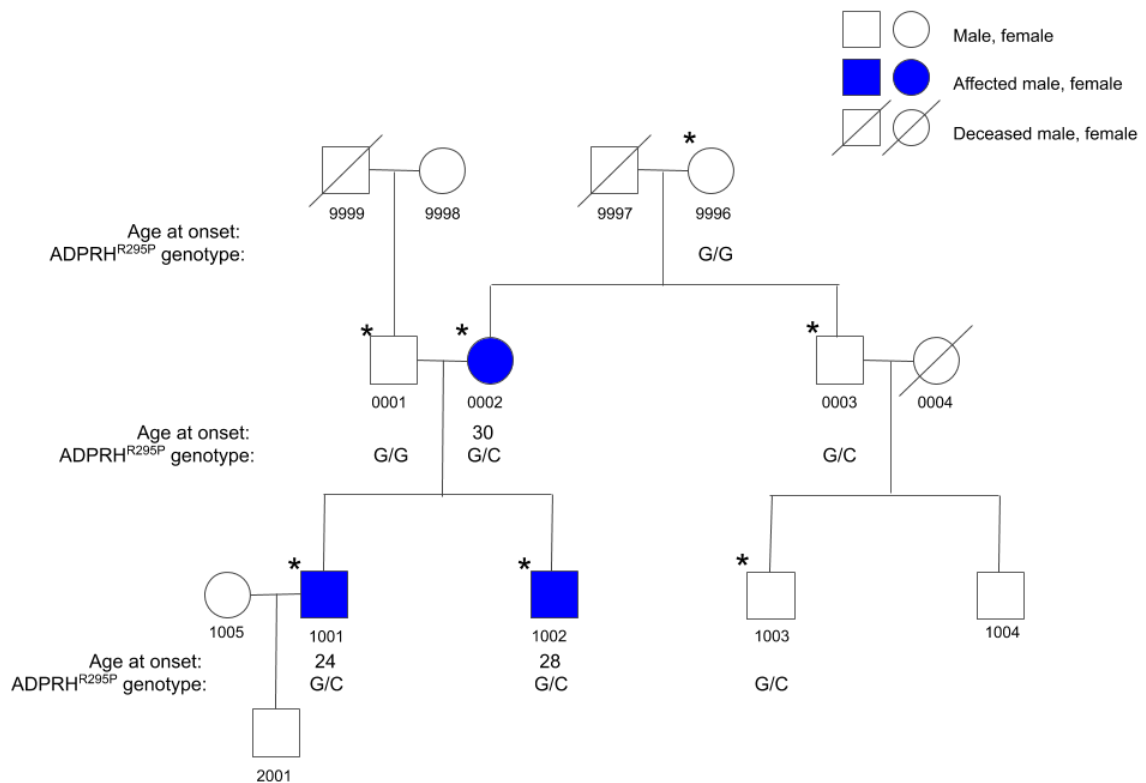


Figure 4.9. UGM471 family tree with the three PLS affected members denoted by coloured in shapes. Genotypes for *ADPRH*^{R295P} variant are included.

Individuals whose DNA samples were available are denoted with an asterisk.

4.2.5. Frequency of the *ADPRH*^{R295P} variant in control population

As the variant *ADPRH*^{R295P} has not been previously described, it cannot be found in variant databases. The only population frequency assessment was done by ASPCR, using

the samples of the control population available in the laboratory. The data is summarised in table 4.18.

Table 4.18. Number of cases of ADPRH^{R295P} variant in different groups.

Group	ADPRH p.R295P G>C		
	GG	GC	CC
Spanish population control	196	0	0
Unaffected family members	2	2	0
Affected family members	0	3	0

A very low frequency of variants in position 295 of ADPRH was observed (Table 4.17). The frequency of the novel variant ADPRH^{R295P} could only be studied in a small control population of 196 samples. Yet, it was not found in any of the control samples.

4.2.6. ADPRH^{R295P} variant's *in silico* analysis

4.2.6.1. ConSurf estimation of evolutionary conservation

Estimation of evolutionary conservation using the ConSurf web server allowed to determine the arginine at position 295 as conserved. Multiple sequence alignment of 150 sequences from different species was performed. 148 of those could be aligned at the R295 position. The only other amino acids known to be present in this position in other species were: histidine, tryptophan, and glutamine. On a scale of 1 (variable) to 9 (conserved) given by ConSurf, the R295 amino acid had an estimated evolutionary conservation score of 8. A multiple sequence alignment with hierarchical clustering was done using MultAlin (Figure 4.10).

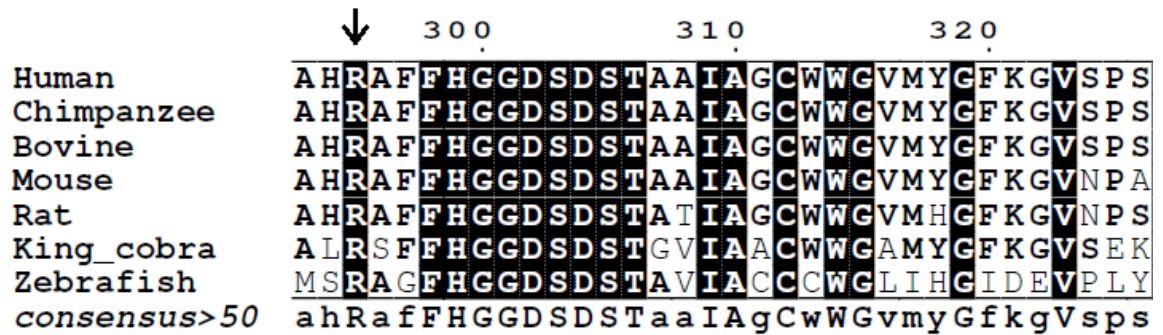


Figure 4.10. Multiple sequence alignment of a segment of the ADPRH protein from seven different species created using MultAlin (Corpet, 1988; <http://multalin.toulouse.inra.fr/multalin/>).

The R295 residue is marked with an arrow. Consensus>50 – threshold between low and high similarity. If residues in a column are identical they have a black background and a capital letter in the consensus row. If they are not, they have a white background and a minuscule letter in the consensus row.

4.2.6.2. HOPE protein structure analysis of ADPRH^{R295P} variant

A general protein structure analysis of the novel variant in ADPRH was performed using HOPE web service. Proline is a smaller residue than arginine. It also has a neutral charge and is more hydrophobic than the positively charged arginine. HOPE web service, using the PDB deposited structure of the protein ADPRH (Figure 4.11; PDB ID: 6IUX; Liu and Yu, 2019), evaluated the contacts of the residue R295. This arginine creates salt bridges with aspartic acid at position 280 and glutamic acid at position 291. A change to proline is predicted to disturb these interactions due to a change in charge.

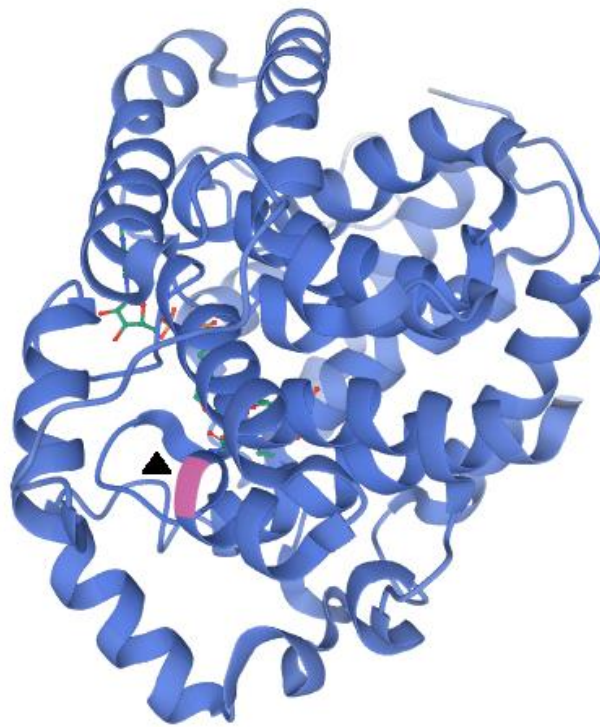


Figure 4.11. Crystal structure of ADPRH protein created using Mol* (Sehna *et al.*, 2017) as deposited in the RCSB Protein Data Bank (Berman *et al.*, 2000; rcsb.org) by Liu and Yu (PDB ID: 6IUX; 2019).

Arginine at position 295 is marked with an arrow and coloured violet. ADP-ribose is visible, marked in red and green.

As seen in figure 4.11, arginine in position 295 is located in an α -helix. In this case, a change to proline is predicted to disturb the α -helix. This can be better seen in a close up of the residue with side chains shown (Figure 4.12). Although R295 is not considered a key amino acid of the active site (Smith *et al.*, 2016), it is located in a domain important for protein activity. The structure analysis predicts this may influence the function and activity of the protein.

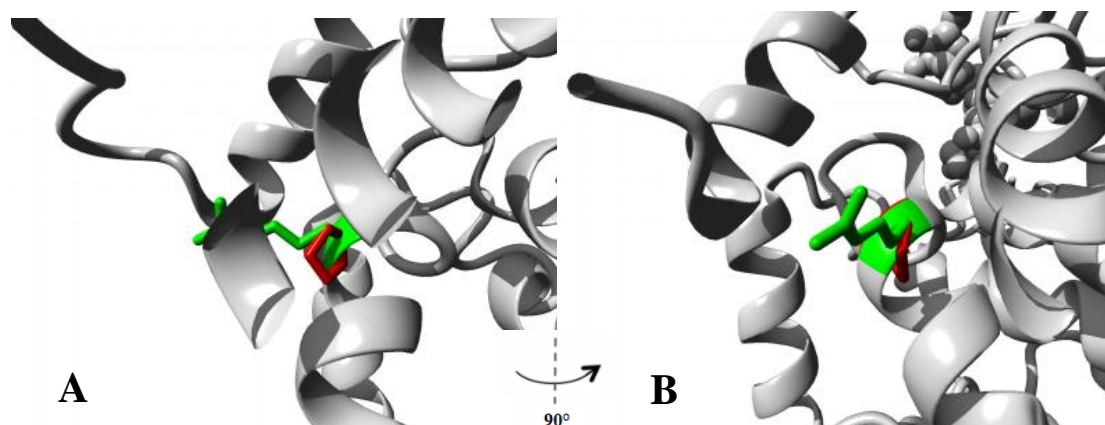


Figure 4.12. Close-up of the residue 295 in ADPRH protein.

The protein is coloured grey, the side chains of both the arginine and the proline residue are shown and coloured green and red respectively. Two different angles are provided: (A) and (B). The images are created by HOPE web service (Venselaar *et al.*, 2010; <https://www3.cmbi.umcn.nl/hope/>), based on a crystal structure of ADPRH protein as deposited in the RCSB Protein Data Bank (Berman *et al.*, 2000; rcsb.org) by Liu and Yu (PDB ID: 6IUX; 2019).

4.2.6.3. Prediction of protein stability by I-Mutant 2.0 web server

I-Mutant 2.0 web server allowed to predict protein stability upon amino acid change. With a reliability index (RI) of 7/10 the stability of ADPRH protein was predicted to decrease upon arginine replacement with proline at position 295. Any amino acid change at that position was predicted to destabilise the protein (Table 4.19), yet the free energy change value (DDG) was the second lowest for arginine to proline change. Only arginine to lysine change was predicted to be more destabilising.

Table 4.19. I-Mutant 2.0 protein stability predictions for SNVs at position R295 in ADPRH.

Position	Amino acid variant	Stability	RI	DDG (kcal/mol)
295	R>V	Decreased	6	-1.38
295	R>L	Decreased	7	-0.94
295	R>I	Decreased	3	-0.87
295	R>M	Decreased	5	-1.16
295	R>F	Decreased	7	-1.04
295	R>W	Decreased	7	-1.01
295	R>Y	Decreased	5	-0.68
295	R>G	Decreased	7	-1.06
295	R>A	Decreased	8	-1.07
295	R>P	Decreased	7	-2.27
295	R>S	Decreased	8	-2.00
295	R>T	Decreased	7	-1.29
295	R>C	Decreased	5	-1.14
295	R>H	Decreased	7	-1.10
295	R>K	Decreased	8	-2.57
295	R>Q	Decreased	8	-1.12
295	R>E	Decreased	8	-0.86
295	R>N	Decreased	6	-1.52
295	R>D	Decreased	8	-1.43

Known existing variants and the novel variant are highlighted in bold text. RI – reliability index ($0 < \text{RI} < 10$); DDG – free energy change value in kcal/mol.

4.2.7. Reassessment of WES results

In light of the discovery of a novel variant ADPRH^{R295P}, its strong destabilizing effect on the protein and presence in two of the unaffected members of the family, a reassessment of the WES results was done. Other variants were evaluated, taking under consideration their interactions with ADPRH or functions in related pathways. The variants found are summarised in table 4.20.

Table 4.20. WES results filtered for possibly damaging variants in genes associated with ADPRH pathways for samples 0002, 1001 and 1002.

Chr	Pos	Ref	Alt	Zigosity	Effect	Gene	Amino acid variant	SIFT prediction (score)	Polyphen prediction (score)
chr13	114,078,558	A	C	HET	missense variant	<i>ADPRHL1</i>	p.L294R	D (0.01)	D (1.0)
chr12	102,559,642	G	C	HET	missense variant	<i>PARPBP</i>	p.G268R	D (0.01)	D (1.0)
chr10	69,644,558	T	C	HET	missense variant	<i>SIRT1</i>	p.S27P	T (1.0)	B (0.0)

Chr – chromosome; Pos – position; Ref – reference allele; Alt – alternate allele; HET – heterozygous; D – deleterious (SIFT)/damaging (Polyphen); T – tolerated; B – benign

The variant *ADPRHL1*^{L294R} can be found in the dbSNP142 and has the accession number rs41306688. *ADPRHL1* is a paralog of ADPRH and rs41306688 has a population frequency of 0.02257 (6351/281400) according to gnomAD (Karczewski *et al.*, 2020; <https://gnomad.broadinstitute.org/>). *PARPBP* encodes the PARP1 Binding Protein. PARP1 is an ADP-ribosyltransferase involved in PARylation. The variant *PARPBP*^{G268R} has not been previously described. *SIRT1* is a part of the sirtuins family. It is an NAD-dependent protein deacetylase with possible MARylation activity.

4.2.7.1. Sanger sequencing verification of variants selected in WES results reassessment

Verification of the WES results by Sanger sequencing is important to eliminate false positive results. Also the manner in which the variants segregate in the family is relevant for classifying their importance for the disease development. The results are summarised in table 4.21.

Table 4.21. Verification of selected variants through Sanger sequencing in family UGM471.

Gene	Amino acid variant	Ref	Alt	Sample ID and diagnosis						
				9996	0001	0002	0003	1001	1002	1003
				non-PLS	non-PLS	PLS	non-PLS	PLS	PLS	non-PLS
<i>ADPRHL1</i>	p.L294R	A	C	A/A	A/A	A/C	A/A	A/C	A/C	A/A
<i>PARPBP</i>	p.G268R	G	C	G/C	G/G	G/G	G/C	G/C	G/C	G/C
<i>SIRT1</i>	p.S27P	T	C	-	-	T/T	-	T/T	T/T	-

Ref – reference allele; Alt – alternate allele

ADPRHL1^{L294R} (rs41306688) segregates with the disease in the family. *PARPBP*^{G268R} not only does not segregate with the disease but also shows a false positive result for sample 0002. The variant *SIRT1*^{S27P} was verified only in the patients' samples, as it showed all false positive results from the WES.

4.2.8. Role of variant *ADPRH*^{R295P} in protein activity

An assay for *ADPRH* activity was devised based on a previously established assay for ADP-ribosyltransferase activity of CT with agmatine as its substrate (Suryadi and Shine, 2011). ADP-ribosylation of agmatine was performed using CT and hydrolysis using *ADPRH*. The *ADPRH* activity was measured in terms of the amount of remaining ADP-ribosylated agmatine after the assay, by use of UPLC. Since ADP-ribosylation of agmatine may occur spontaneously, a control was used where neither CT nor *ADPRH* were added. To control ADP-ribosylation by CT, a sample was prepared where CT but not *ADPRH* was added.

The experiment was performed by cleavage of ADP-ribose from agmatine using *ADPRH*^{WT}, *ADPRH*^{R295P} and *ADPRH*^{D55A/D56A} variant lacking cleavage activity. *ADPRH*^{D55A/D56A} is a variant created in the laboratory as residues D55 and D56 have been described as key residues of the active site of the protein (Smith *et al.*, 2016). Since the reaction mechanism of *ADPRH* relies on magnesium cations coordinated by aspartic acid residues (D55, D56, D302, D304; Mueller-Dieckmann *et al.*, 2006; Berthold *et al.*, 2009), introducing changes p.D55A/D56A should alter the active site and the protein activity

significantly. The activity of ADPRH was, therefore, compared to that of the novel variant and the variant with active site of the protein affected.

First the reaction products needed to be identified by separation using the UPLC data from a detector measuring absorbance at 260 nm. The details of the reaction protocol can be found in the methods section. In order to identify the peak for ADP-ribosyl-agmatine on the chromatogram, several control runs were made (Table 4.22).

Table 4.22. Control runs of the ADPRH activity assay.

Run	Reaction reactants and reagents				
	Agmatine	β -NAD ⁺	DTT	sodium phosphate	CT
1.	-	+	-	-	-
2.	+	-	-	-	-
3.	+	+	+	+	-
4.	-	+	+	+	+
5.	+	+	+	+	+

“+” – added to the reaction and run on UPLC; “-“ – not added to the reaction and run on UPLC

These runs allowed for identification of the peak for ADP-ribosyl-agmatine close to retention time (RT) 1.45 minutes (Figure 4.13).

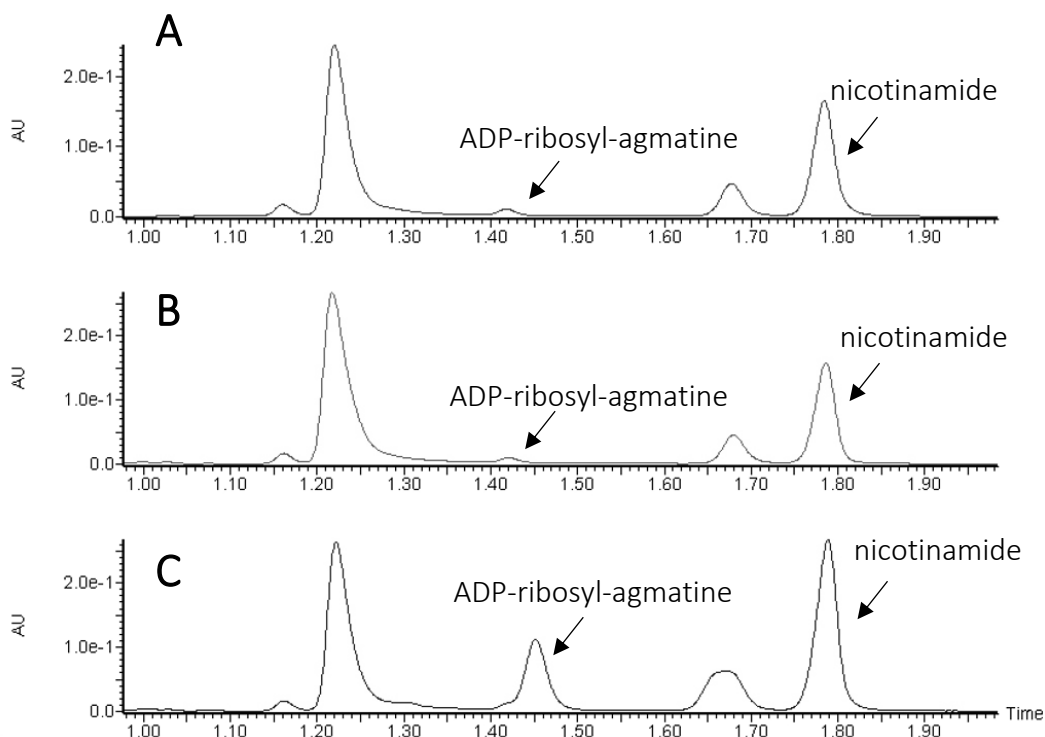


Figure 4.13. UPLC chromatograms representing three separations to identify ADP-ribosyl-agmatine.

(A) Run 3 of table 4.22. All reactants, no CT. (B) Run 4. All reactants except agmatine, with CT. (C) Run 5. All the reactants, with CT. ADP-ribosyl-agmatine appeared at retention time (RT) 1.45 minutes in the run 5. Data obtained from three technical and biological replicates.

Experimental runs were performed from products of reactions with ADPRH and its variants. Appropriate controls were used (Table 4.23).

Table 4.23. Experimental runs of the ADPRH activity assay.

Run	Reaction reactants and reagents							
	Agmatine	β -NAD ⁺	DTT	sodium phosphate	CT	ADPRH ^{WT}	ADPRH ^{R295P}	ADPRH ^{D55A/D56A}
1.	+	+	+	+	+	+	-	-
2.	+	+	+	+	+	-	+	-
3.	+	+	+	+	+	-	-	+
4.	+	+	+	+	+	-	-	-
5.	+	+	+	+	-	-	-	-

+ – added to the reaction and run on UPLC; - – not added to the reaction and run on UPLC

In order to measure the relative activity of ADPRH and its variants, the area under ADP-ribosyl-agmatine peaks was measured for all reactions (Figure 4.14).

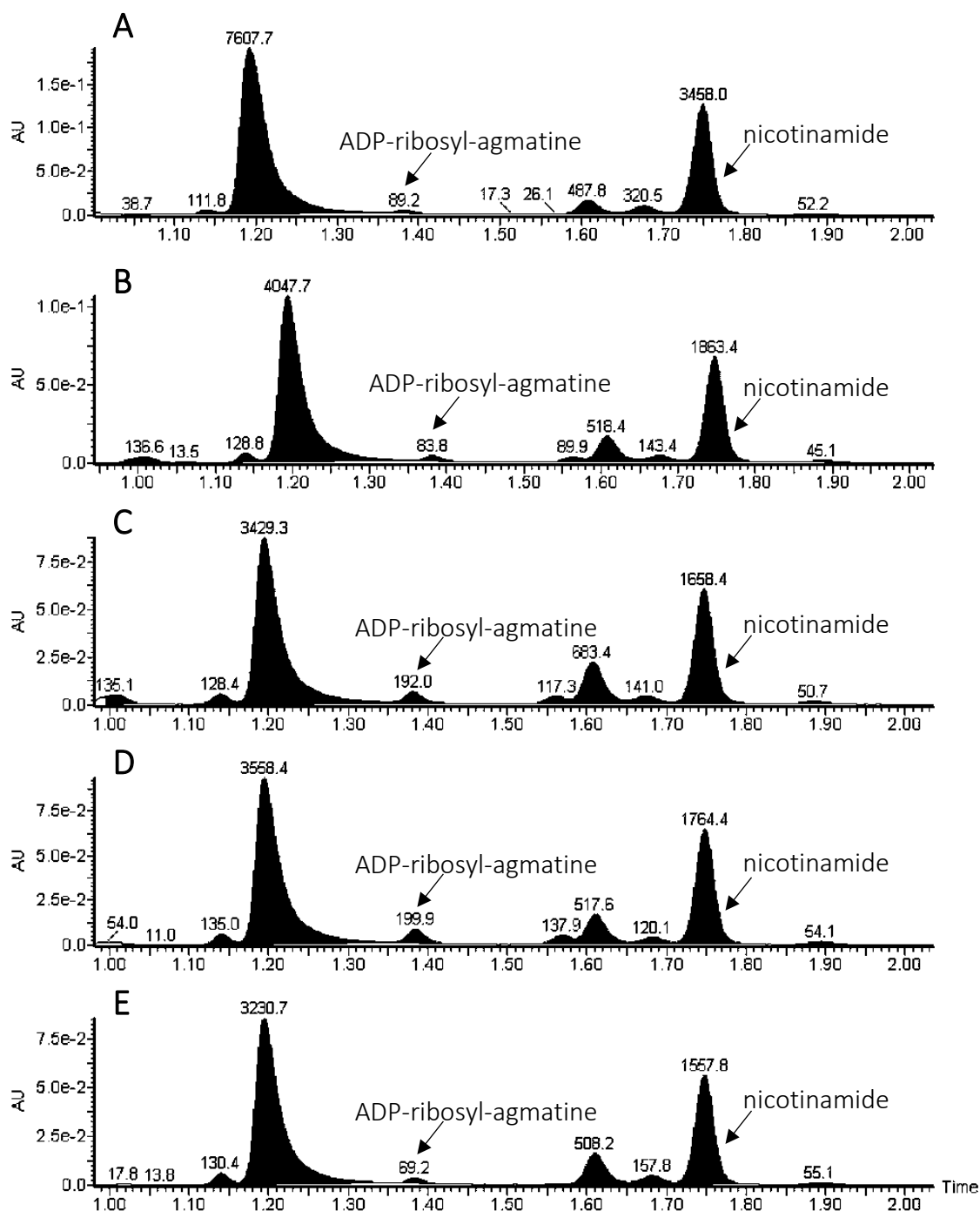


Figure 4.14. Selected UPLC chromatograms representing five separations to measure the relative activity of ADPRH and its variants ADPRH^{R295P} and ADPRH^{D55A/D56A}.

Figure 4.14. Selected UPLC chromatograms representing five separations to measure the relative activity of ADPRH and its variants ADPRH^{R295P} and ADPRH^{D55A/D56A} (continued).

(A) Run 1. Reaction with ADPRH^{WT}, (B) run 2. Reaction with ADPRH^{R295P}, (C) run 3. Reaction with ADPRH^{D55A/D56A}, (D) run 4. Control reaction without ADPRH variants, (E) run 5. Control reaction with all reactants without CT or ADPRH. Data obtained from three technical and biological replicates.

Both ADPRH^{WT} and ADPRH^{R295P} were able to hydrolyse the CT-ADP-ribosylated agmatine, bringing its levels back to those of spontaneous ADP-ribosylation (Absorbance-minute units, mean \pm SEM: 86.2 \pm 2.2 for ADPRH^{WT}, 80.4 \pm 2.5 for ADPRH^{R295P}, 79.1 \pm 12.0 for spontaneous ADP-ribosylation). They both cleaved ADP-ribose with similar efficiency (Figure 4.15). The double variant ADPRH^{D55A/D56A} did not cleave ADP-ribose efficiently (185.7 \pm 3.9 for ADPRH^{D55A/D56A}). ADP-ribosyl-agmatine levels were as high for reaction with ADPRH^{D55A/D56A} as after the reaction with CT (206.4 \pm 5.0 for CT). ADPRH^{D55A/D56A} did not show enzymatic activity and differed significantly from ADPRH^{WT} ($p < 0.001$) and ADPRH^{R295P} ($p < 0.001$).

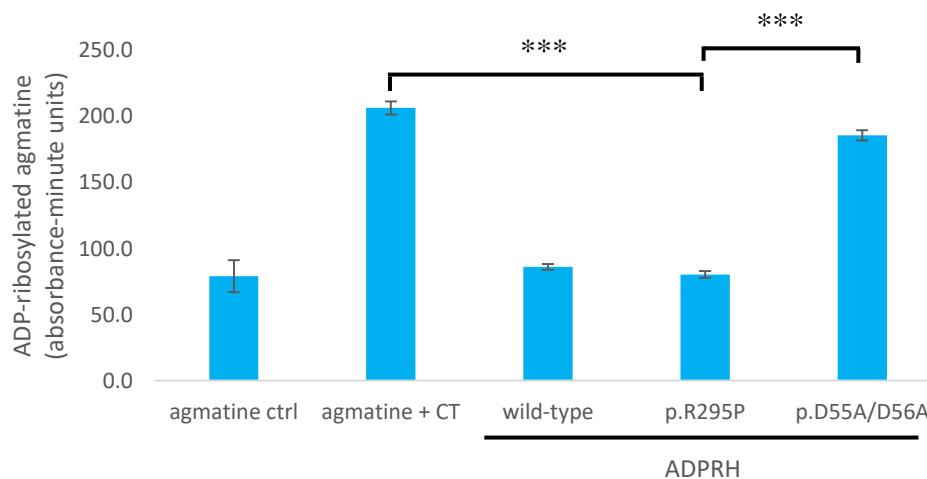


Figure 4.15. ADPRH activity assay with agmatine and CT using UPLC.

Agmatine ctrl – control reaction with all reactants without CT or ADPRH; agmatine + CT – control reaction without ADPRH variants; ADPRH: wild-type – reaction with wild-type ADPRH, p.R295P – reaction with ADPRH^{R295P}, p.D55A/D56A – reaction with ADPRH^{D55A/D56A}. *** $p < 0.001$, one-way ANOVA comparing controls, ADPRH^{WT}, ADPRH^{R295P} and ADPRH^{D55A/D56A}, with *post hoc* Tukey multiple comparisons of means test. Data obtained from three technical and biological replicates.

4.2.9. Role of variant ADPRH^{R295P} in mRNA expression

Involvement of variant ADPRH^{R295P} in mRNA expression was studied by the use of RT-qPCR. Previously described variants: rs150984649 (ADPRH^{R295Q}) and rs190409478 (ADPRH^{R295*}), at the same position as the novel ADPRH^{R295P}, were used as controls. The variant ADPRH^{R295Q} had a change of glutamine for arginine which was predicted to be more stable than that introduced by ADPRH^{R295P} (Table 4.19). ADPRH^{R295*} encoded for a truncated protein with a premature stop codon. The double variant ADPRH^{D55A/D56A} and ADPRH^{WT} were also used to determine ADPRH^{R295P} effect on mRNA levels. Transfection efficiency was controlled by amplification of kanamycin from the plasmid resistance cassette.

Although mRNA levels of ADPRH^{R295P}, ADPRH^{R295Q} and ADPRH^{D55A/D56A} were slightly lower than those of ADPRH^{WT} (Normalised mean amplification efficiency \pm SEM: 1.000 \pm 0.115 for ADPRH^{WT}; 0.713 \pm 0.057 for ADPRH^{R295P}; 0.574 \pm 0.091 for ADPRH^{R295Q}; 0.814 \pm 0.154 for ADPRH^{D55A/D56A}), the difference was not significant (Figure 4.16). Only the mRNA levels of ADPRH^{R295*} truncating variant were significantly lower ($p < 0.001$ for ADPRH^{WT} and ADPRH^{D55A/D56A}; $p = 0.002$ for ADPRH^{R295P}; $p = 0.036$ for ADPRH^{R295Q}) at mean amplification efficiency of 0.108 \pm 0.013 SEM.

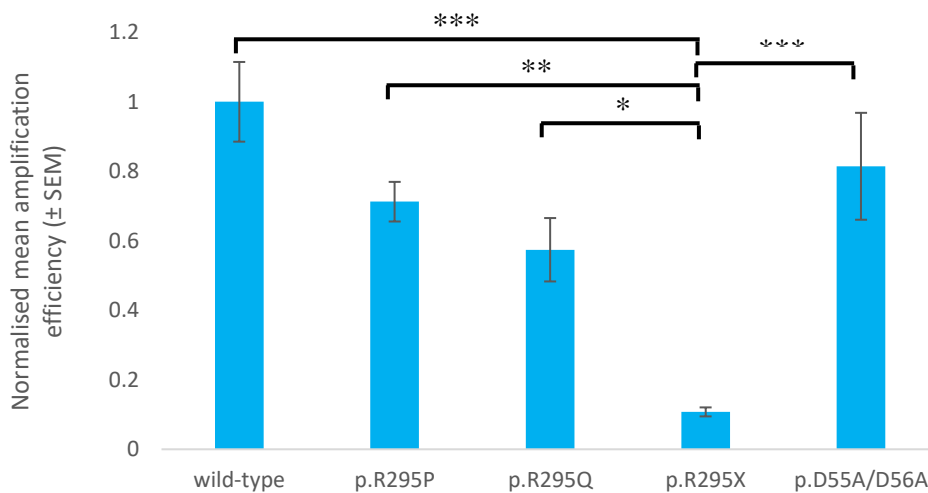


Figure 4.16. RT-qPCR results for mRNA levels of ADPRH^{WT}, ADPRH^{R295P}, ADPRH^{R295Q}, ADPRH^{R295*} and ADPRH^{D55A/D56A} overexpressed in SH-SY5Y cells.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ one-way ANOVA comparing ADPRH^{WT} and its variants, with *post hoc* Tukey multiple comparisons of means test. Data obtained from three technical and two biological replicates.

Regardless of missense variant studied, the mRNA levels did not change in respect to the ADPRH^{WT}, unlike the nonsense variant whose mRNA levels were lower.

4.2.10. Role of variant ADPRH^{R295P} in protein expression

Analysis of protein overexpression using the plasmids carrying the ADPRH^{WT} gene and the ADPRH^{R295P}, ADPRH^{R295Q}, ADPRH^{R295*} and ADPRH^{D55A/D56A} variants was done using WB and SH-SY5Y cells. The protein levels of ADPRH^{R295P} (Normalised mean band intensity \pm SEM: 0.082 ± 0.011) were significantly lower than those of ADPRH^{WT} (1.000 ± 0.114 ; $p = 0.005$), ADPRH^{R295Q} (0.768 ± 0.242 ; $p = 0.034$) and ADPRH^{D55A/D56A} (1.132 ± 0.149 ; $p = 0.002$; Figure 4.17).

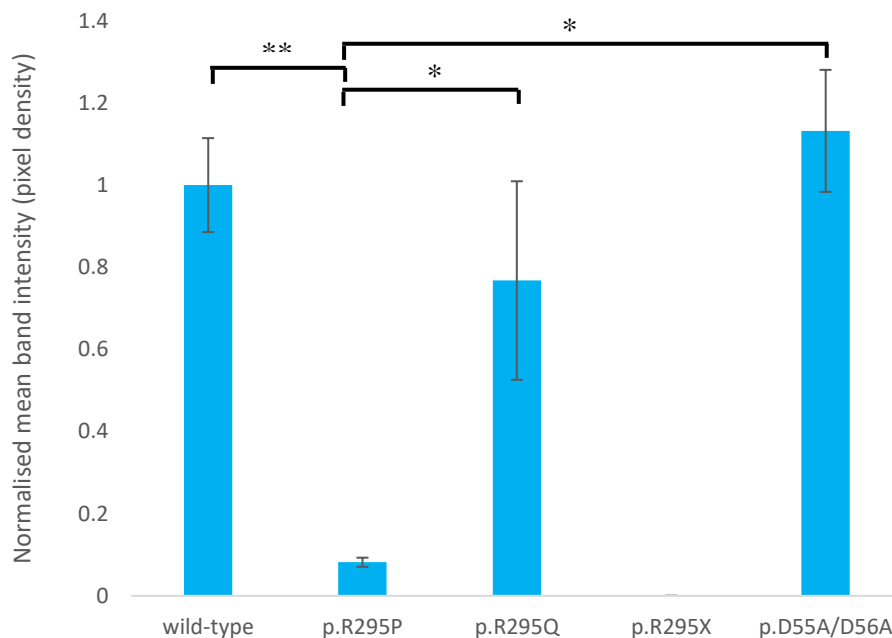


Figure 4.17. Western blot results for ADPRH^{WT} and ADPRH^{R295P}, ADPRH^{R295Q}, ADPRH^{R295*} and ADPRH^{D55A/D56A} variants overexpression in SH-SY5Y cells.

* $p < 0.05$, ** $p < 0.01$, one-way ANOVA comparing ADPRH^{WT} and its variants, with *post hoc* Tukey multiple comparisons of means test. Data obtained from three technical and biological replicates.

There were no differences in protein levels between the ADPRH^{WT} and ADPRH^{R295Q} and ADPRH^{D55A/D56A} variants. The variant expressing the truncated protein ADPRH^{R295*} was not present in the sample, showing no band on the membrane.

4.2.11. Protein-protein interaction between ADPRH and ADPRHL1

Co-IP was performed with the wild-type forms of ADPRH and ADPRHL1. Data was obtained from two technical and biological replicates. While the proteins precipitated correctly, they did not co-precipitate. The co-IP gave a negative result and no proof of interaction between wild-type forms of ADPRH and ADPRHL1.

4.2.12. PSEN1^{L166P} (rs63750265) mutant and AD

In 2019, Jazi *et al.* have associated mutations PSEN1^{A431E} and PSEN1^{L381V} with PLS. In 2020 Vazquez-Costa (forthcoming) has identified a mutation PSEN1^{L166P} with the accession number rs63750265 in the PLS patients from family UGM471. PSEN1^{L166P} is a rare mutation associated with a very early onset of AD symptoms and spastic paraparesis (Moehlmann *et al.*, 2002; Lyoo *et al.*, 2016). The WES result for PSEN1^{L166P} is summarised in table 4.24.

Table 4.24. WES results for PSEN1^{L166P} (rs63750265) mutation for samples 0002, 1001 and 1002.

Chr	Pos	Ref	Alt	Zigosity	Effect	Gene	Amino acid variant	SIFT prediction (score)	Polyphen prediction (score)
chr14	73,653,577	T	C	HET	missense mutation	PSEN1	p.L166P	D (0.01)	D (0.981)

Chr – chromosome; Pos – position; Ref – reference allele; Alt – alternate allele; HET – heterozygous; D – deleterious (SIFT)/damaging (Polyphen)

Several other mutations are recorded at position L166 of PSEN1 which accompany a clinical phenotype of AD. There is PSEN1^{L166del} (rs63751458; Knight *et al.*, 2007), PSEN1^{L166H} (rs63750265; Pantieri *et al.*, 2005), PSEN1^{L166R} (rs63750265; Ezquerra *et al.*, 2000) and PSEN1^{L166V} (Sassi *et al.*, 2014).

4.2.12.1. Sanger sequencing verification of mutation PSEN1^{L166P} (rs63750265)

Sanger sequencing was performed with samples from all available UGM471 family members to confirm lack of false positive WES results and possible segregation of the variant with the disease (Table 4.25).

Table 4.25. Verification of mutation PSEN1^{L166P} (rs63750265) through Sanger sequencing in family UGM471.

Gene	Amino acid variant	Ref	Alt	Sample ID and diagnosis						
				9996	0001	0002	0003	1001	1002	1003
<i>PSEN1</i>	p.L166P	T	C	T/T	T/T	T/C	T/T	T/C	T/C	T/T

Ref – reference allele; Alt – alternate allele

The mutation segregated with the disease as summarised in figure 4.18.

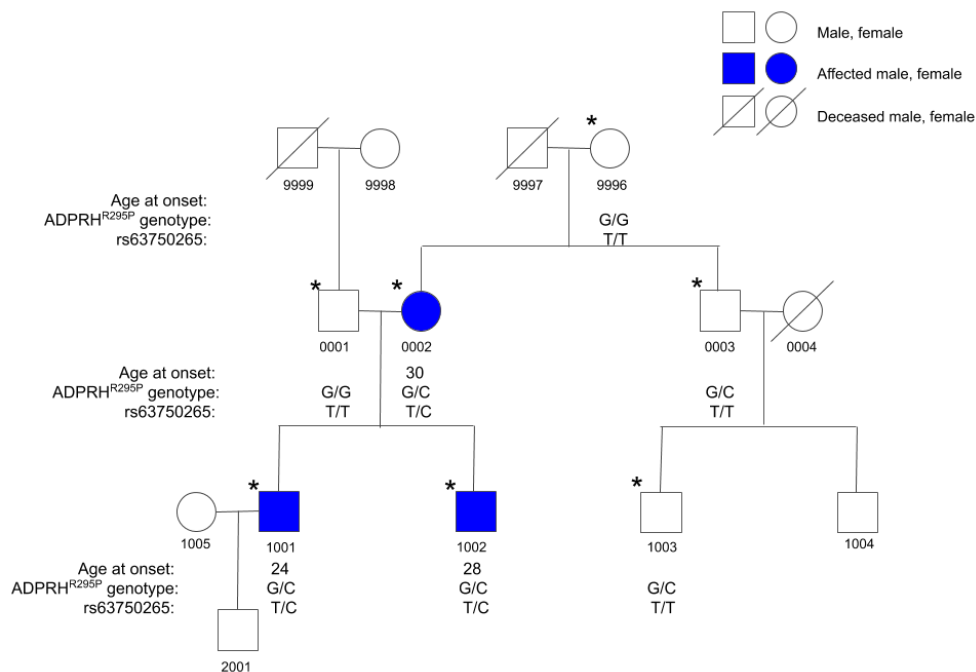


Figure 4.18. UGM471 family tree with the three PLS affected members denoted by coloured in shapes. Genotypes for PSEN1^{R295P} (rs63750265) mutant and ADPRH^{R295P} variant are included.

Individuals whose DNA samples were available are denoted with an asterisk.

4.3. Study of etiopathogenesis of PKD in family UGM478

4.3.1. Clinical data

Five patients from family UGM478 were diagnosed with PKD. The proband was examined with neuroimaging and EEG and others were diagnosed through clinical history and neurological examination. The results of the patients' clinical data are summarised in table 4.26.

Table 4.26. Family UGM478 patient data.

Patient	Sex	Age of onset (years)	Attack duration (seconds)	Attack frequency	Aura	Triggers	Treatment	Disease evolution
1003	F	14	< 10	< 1/month	No	Sudden movement, stress	None	Remission at 23 years of age
1005	M	10	< 10	10-20/day	No	Sudden movement, stress	Yes, unknown	Remission at 25 years of age
2004	M	12	< 10	< 1/week	Yes	Sudden movement	None	Remission at 15 years of age
2005	M	14	< 10	> 20/day	Yes	Sudden movement, stress, fatigue, shock	CBZ	Withdrawing CBZ and without attacks
2007	M	8	< 10	1-10/day	Yes	Sudden movement	CBZ	Without attacks with CBZ treatment

M – male; F – female; CBZ – Carbamazepine.

4.3.2. WES data

4.3.2.1. Base-calling and image analysis raw data

Basic raw data obtained through base-calling and image analysis produced by WES for samples 1005, 2005 and 2007 are summarised in table 4.27.

Table 4.27. WES raw data: reads, base content and phred quality scores for samples 1005, 2005 and 2007.

Sample ID	Total read bases (bp)	Total reads	GC (%)	Q20 (%)	Q30 (%)
1005	9,094,496,318	90,044,518	49.5	98.8	96.6
2005	9,321,758,236	92,294,636	49.5	98.7	96.3
2007	8,105,274,644	80,250,244	49.6	98.7	96.3

Total read bases – Total number of bases sequenced; Total reads – Total number of reads. In Illumina paired-end sequencing, read1 and read2 are added; GC (%) – GC content; Q20 (%) – Ratio of reads that have phred quality score of over 20; Q30 (%) – Ratio of reads that have phred quality score of over 30.

4.3.2.2. Read alignment results

The average read length for all samples was 101 bp. The post-alignment statistics are given in table 4.28.

Table 4.28. WES post-alignment statistics for samples 1005, 2005 and 2007.

Sample ID	1005	2005	2007
Initial Mappable Reads	90,013,269	92,258,890	80,218,518
% Initial Mappable Reads	99.9	99.9	99.9
Non-Redundant Reads	78,707,862	81,119,190	71,692,113
% Non-Redundant Reads	87.4	87.9	89.3
On-Target Reads	65,943,619	67,267,032	59,662,592
% On-Target Reads	83.7	82.9	83.2
On-Target Yield (bp)	5,782,503,371	5,892,020,939	5,223,500,863
Mean Depth of Target Regions (X)	114.7	116.9	103.6

Table 4.28. WES post-alignment statistics for samples 1005, 2005 and 2007 (continued).

Initial Mappable Reads – Number of mapped reads to human genome; % Initial Mappable Reads – $100 * (\text{Initial mappable reads}) / (\text{Total reads})$; Non-Redundant Reads – Number of de-duplicate reads from Picard tools; % Non-Redundant Reads – $100 * (\text{Non-redundant reads}) / (\text{Initial mappable reads})$; On-Target Reads – Number of reads mapped to target regions; % On-Target Reads – $100 * (\text{On-target reads}) / (\text{Non-redundant reads})$; On-Target Yield (bp) – The sum of the bases in the final alignment to the target regions; Mean Depth of Target Regions (X) – $(\text{On-target yield}) / (\text{Target regions})$

4.3.2.3. SNP calling results

SNP calling allowed for identification of variants in DNA sample sequences. The SNP calling results for samples 1005, 2005 and 2007 are summarised in table 4.29.

Table 4.29. WES SNP calling summary for samples 1005, 2005 and 2007.

Sample ID	1005	2005	2007
Number of SNPs	78,467	80,074	78,579
Synonymous Variants	11,467	11,613	11,472
Missense Variants	10,431	10,513	10,526
Stop Gained	90	83	79
Stop Lost	35	35	39
Number of INDELs	9,151	9,498	9,132
Frameshift Variants	260	267	282
Inframe Insertions	144	158	154
Inframe Deletions	186	191	190
% Found in dbSNP142	97.5	97.4	97.5
Het/Hom Ratio	1.4	1.5	1.5
Ts/Tv Ratio	2.3	2.3	2.3

Table 4.29. WES SNP calling summary for samples 1005, 2005 and 2007 (continued).

% Found in dbSNP142 – percentage of Number of SNPs found in dbSNP142; Het/Hom Ratio – Ratio of number of heterozygous variants to number of homozygous variants; Ts/Tv Ratio – Ratio of transition rate of SNVs that pass the quality filters divided by transversion rate of SNVs that pass the quality filters. Transitions are interchanges of purines (A,G) or of pyrimidines (C, T). Transversions are interchanges between purine and pyrimidine bases.

4.3.3. WES results filtration and identification of interesting variants

The results obtained from WES analysis were filtered and the most interesting variants assessed. Firstly, genes known to be associated with PKD were looked at, to see whether a known or a new variant in those genes could cause or affect the disease (Table 4.30).

Table 4.30. WES results filtered for PKD and PxD associated genes for samples 1005, 2005 and 2007.

Gene	Number of intron variants	Number of synonymous variants	Number of missense variants	Other variants
<i>PRRT2</i>	0	1	0	2 upstream gene variants 3'-UTR variant stop gain variant
<i>MR-1</i>	1	0	0	3'-UTR variant
<i>ATP1A3</i>	0	1	0	3'-UTR variant
<i>SLC2A1</i>	2	0	0	0
<i>ECHS1</i>	2	1	2	0
<i>PDHX</i>	5	2	1	2 5'-UTR variants
<i>DLAT</i>	2	0	0	0

Of the variants in table 4.30 only one had a putative impact predicted as HIGH by SnpEff software. It was the novel variant *PRRT2*^{Q106*}. It was not previously described in any variant database. It encodes a PTC c.C316T p.Q106*. Other variants in table 4.30 were

filtered out from further analysis either by prediction of their impact, their frequency in population or a review of current knowledge on their effect.

All variants, whether related to PKD or not, went through a restrictive filtering process to assess the possible cause of PKD in family UGM478. Filtering was done as described in the methods section. An important step was the filtering by role, function, tissue of expression and pathways the variants are involved in (Table 4.31).

No interesting nonsense variant was found other than PRRT2^{Q106*}. *ARHGEF10L* encodes a protein involved in signal transduction, *DLGAP2* encodes a protein which may be important for synapse organization and signalling in neurons, *ZPRI* encodes a protein involved in neuronal differentiation, axonal growth and spinal muscular atrophy, *RYR1* is involved in mediating Ca²⁺ release in neurons, which may stimulate prolonged signalling in the brain, *MED25* mutations are associated with Charcot-Marie-Tooth disease also known as the hereditary motor and sensory neuropathy (Stelzer *et al.*, 2016; www.genecards.org), *Sorbs2* knock-out mice had impaired dendritic development (Zhang *et al.*, 2016), *SIN3A* encodes a protein involved in neurodevelopmental disorders (Witteveen *et al.*, 2016). Excluding DLGAP2^{R37W} (rs988565981), all of the variants from table 4.31 are novel.

The standards and guidelines for the interpretation of sequence variants (Richards *et al.*, 2015) and the gene constraint observed/expected score (Karczewski *et al.*, 2020; <https://gnomad.broadinstitute.org/>) were used to assess these variants (Table 4.32). Low values of the observed/expected score are indicative of strong intolerance.

Table 4.31. WES results filtered for possibly damaging variants for samples 1005, 2005 and 2007.

Chr	Pos	Ref	Alt	Zigosity	Gene	Amino acid variant	SIFT prediction (score)	Polyphen prediction (score)	Putative Impact (SnEff)
chr1	17,965,061	T	C	HET	<i>ARHGEF10L</i>	p.L683P	D (0.0)	D (1.0)	MODERATE
chr4	186,545,342	C	T	HET	<i>SORBS2</i>	p.C510Y	D (0.015)	D (0.998)	MODERATE
chr8	1,496,968	C	T	HET	<i>DLGAP2</i>	p.R37W	D (0.0)	D (1.0)	MODERATE
chr11	116,656,234	A	G	HET	<i>ZPR1</i>	p.L234P	D (0.001)	D (0.992)	MODERATE
chr15	75,682,153	T	A	HET	<i>SIN3A</i>	p.D954V	D (0.003)	D (0.961)	MODERATE
chr16	29,824,691	C	T	HET	<i>PRRT2</i>	p.Q106*	-	-	HIGH
chr19	38,959,720	G	A	HET	<i>RYR1</i>	p.E1166K	D (0.001)	D (0.998)	MODERATE
chr19	50,338,794	G	C	HET	<i>MED25</i>	p.G560R	D (0.001)	D (1.0)	MODERATE

Chr – chromosome; Pos – position; Ref – reference allele; Alt – alternate allele; HET – heterozygous; HOM – homozygous; D – deleterious (SIFT)/damaging (Polyphen). SIFT and Polyphen both do not perform predictions for nonsense variants.

Table 4.32. Interpretation of the selected variants according to Richards *et al.* (2015) and observed/expected score.

Gene	Amino acid variant	Effect	Sequence variant classification	observed/expected score (90 % CI)
<i>ARHGEF10L</i>	p.L683P	missense variant	Uncertain significance	0.85 (0.8 – 0.91)
<i>DLGAP2</i>	p.R37W	missense variant	Uncertain significance	1.11 (1.04 – 1.18)
<i>MED25</i>	p.G560R	missense variant	Uncertain significance	0.83 (0.77 – 0.91)
<i>PRRT2</i>	p.Q106*	stop gain	Pathogenic	0.18 (0.07 – 0.56)
<i>RYR1</i>	p.E1166K	missense variant	Uncertain significance	0.9 (0.87 – 0.93)
<i>SIN3A</i>	p.D954V	missense variant	Uncertain significance	0.54 (0.49 – 0.59)
<i>SORBS2</i>	p.C510Y	missense variant	Uncertain significance	0.99 (0.93 – 1.05)
<i>ZPR1</i>	p.L234P	missense variant	Uncertain significance	0.87 (0.78 – 0.97)

4.3.4. Sanger sequencing verification of selected variants

The variants selected from WES results needed to be confirmed by Sanger sequencing to eliminate a possibility of a false positive result and to see whether they segregate with the disease. As only samples 1005, 2005 and 2007 were used in WES, samples of other family members, including other patients were added to the Sanger sequencing confirmation when necessary (Table 4.33).

Table 4.33. Verification of selected variants through Sanger sequencing in family UGM478.

Gene	Amino acid variant	Ref	Alt	Sample ID and diagnosis						
				1003 PKD	1005 PKD	1006 non- PKD	2004 PKD	2005 PKD	2006 non- PKD	2007 PKD
<i>ARHGEF10L</i>	p.L683P	T	C	-	T/C	-	-	T/T	-	T/C
<i>DLGAP2</i>	p.R37W	C	T	-	C/T	-	-	C/C	-	C/T
<i>MED25</i>	p.G560R	G	C	-	G/C	-	-	G/G	-	G/C
<i>PRRT2</i>	p.Q106*	C	T	C/T	C/T	C/C	C/T	C/T	C/C	C/T
<i>RYR1</i>	p.E1166K	G	A	-	G/A	-	-	G/A	-	G/G
<i>SIN3A</i>	p.D954V	T	A	-	T/A	-	-	T/T	-	T/A
<i>SORBS2</i>	p.C510Y	C	T	-	C/T	-	-	C/C	-	C/T
<i>ZPR1</i>	p.L234P	A	G	-	A/G	-	-	A/A	-	A/G

Ref – reference allele; Alt – alternate allele; “-“ – not sequenced

Sanger sequencing confirms the WES results for PRRT2^{Q106*} novel variant, and shows its segregation with the disease. No other variant from table 4.33 segregates with the disease in family UGM478. Figure 4.19 shows the UGM478 family tree with the genotype of each member whose DNA sample was available for PRRT2^{Q106*} variant.

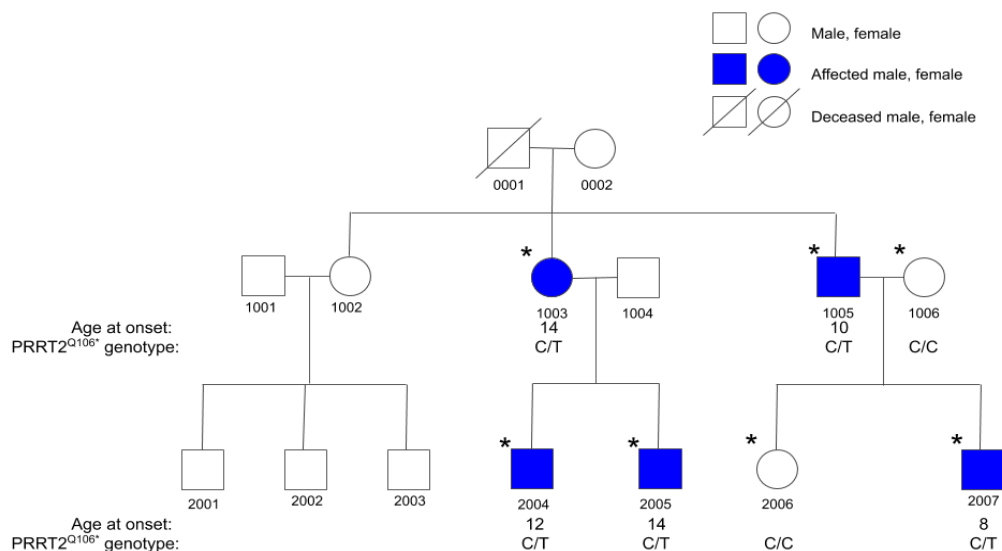


Figure 4.19. UGM478 family tree with the five PKD affected members denoted by coloured in shapes. Genotypes for PRRT2^{Q106*} variant are included.

Individuals whose DNA samples were available are denoted with an asterisk.

4.3.5. Frequency of the PRRT2^{Q106*} variant in control population

As the variant PRRT2^{Q106*} is novel, it cannot be found in variant databases. The only population frequency assessment was done by ASPCR, using the samples of the control population available in the laboratory. The data is summarised in table 4.34.

Table 4.34. Number of cases of PRRT2^{Q106*} variant in different groups.

Group	PRRT2 p.Q106X C>T		
	CC	CT	TT
Spanish population control	192	0	0
Unaffected family members	2	0	0
Affected family members	0	5	0

Due to the lack of information on population frequency of variant PRRT2^{Q106*} the frequency could only be studied in a small control population of 192 samples. The variant was not found in any of the control samples. It was found in heterozygosis in all 5 PKD patients from family UGM478 and was absent in the healthy members of the family.

4.3.6. Role of variant PRRT2^{Q106*} in mRNA expression

Truncating mutations in *PRRT2* are a common cause of PKD (Valtorta *et al.*, 2016). Wu *et al.* (2014) suggest the PTC in *PRRT2* trigger the NMD pathway. Previously described truncating variants PRRT2^{Q163*} (Wang *et al.*, 2011) and PRRT2^{Q250*} (Ono *et al.*, 2012) were used here as controls together with the PRRT2^{WT}. PRRT2^{Q106*}, PRRT2^{Q163*} and PRRT2^{Q250*} encode a PTC replacing a glutamine at different domains of the protein. The novel variant PRRT2^{Q106*} introduces a stop codon before the proline-rich domain of PRRT2, the variant PRRT2^{Q163*} in the middle of it, and variant PRRT2^{Q250*} after (Figure 4.20). None of these variants encoded for transmembrane domains of the protein. Transfection efficiency was controlled by amplification of kanamycin from the plasmid resistance cassette.

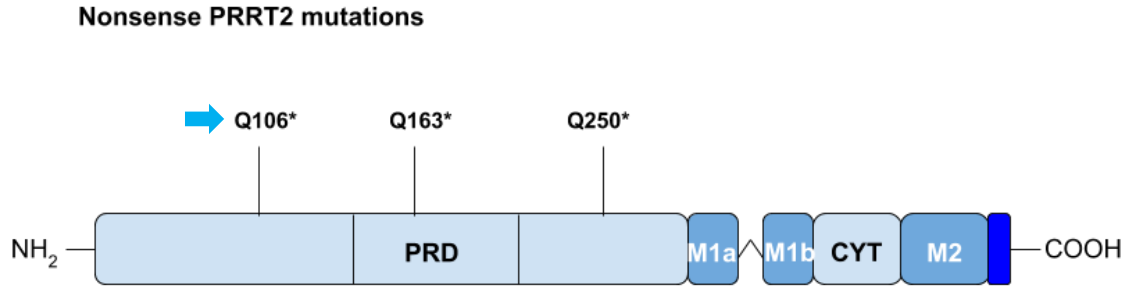


Figure 4.20. Structure of PRRT2 protein with map of nonsense variants PRRT2^{Q106*}, PRRT2^{Q163*} and PRRT2^{Q250*} with regards to their protein domain location (modified with permission from Valtorta *et al.*, 2016).

The novel variant PRRT2^{Q106*} is marked with an arrow. See appendix I for license agreement.

Levels of mRNA were lower for all tested variants (PRRT2^{Q106*}, PRRT2^{Q163*} and PRRT2^{Q250*}; $p < 0.001$) in comparison with the PRRT2^{WT} (Figure 4.21).

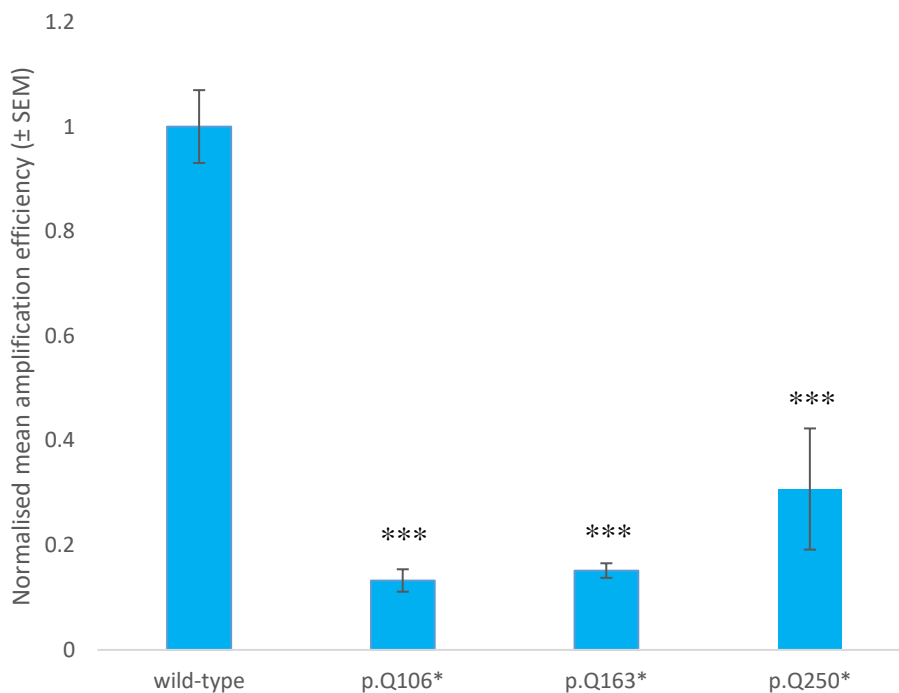


Figure 4.21. RT-qPCR results for mRNA levels of PRRT2^{WT} and PRRT2^{Q106*}, PRRT2^{Q163*} and PRRT2^{Q250*} variants overexpressed in SH-SY5Y cells.

*** $p < 0.001$ one-way ANOVA comparing PRRT2^{WT} and its variants, with *post hoc* Tukey multiple comparisons of means test. Data obtained from three technical and two biological replicates.

The results were normalised to represent the PRRT2^{WT} mRNA levels as 100 %. The normalised mean amplification efficiency \pm SEM was: $1.000 \pm 0,070$ for PRRT2^{WT}; 0.133 ± 0.021 for PRRT2^{Q106*}; 0.152 ± 0.014 for PRRT2^{Q163*}; 0.307 ± 0.116 for PRRT2^{250*}. No significant difference was found between variants.

NMD was inhibited in the NMDI14 treated group to test the theory that PRRT2 mRNA with PTC is decayed through this pathway (Wu *et al.*, 2014). mRNA levels increased for all variants (Figure 4.22). From 0.133 ± 0.021 to 0.588 ± 0.040 for PRRT2^{Q106*} ($p < 0.001$), from 0.152 ± 0.014 to 0.487 ± 0.085 for PRRT2^{Q163*} ($p = 0.011$) and from 0.307 ± 0.116 to 0.590 ± 0.045 for PRRT2^{250*} (not statistically significant at $p = 0.05$).

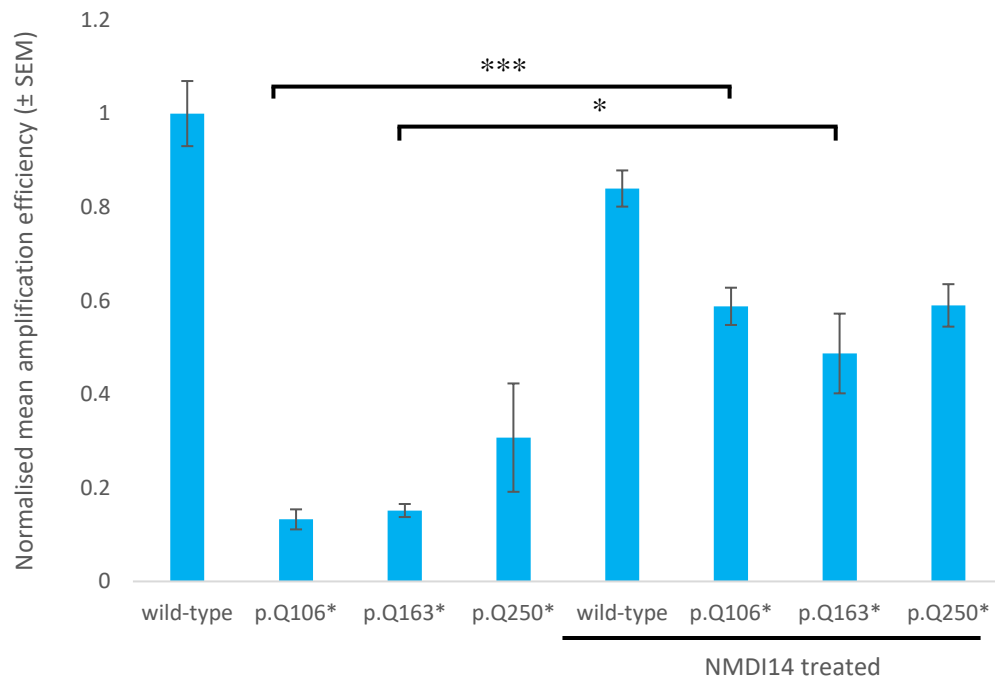


Figure 4.22. RT-qPCR results for mRNA levels of PRRT2^{WT} and PRRT2^{Q106*}, PRRT2^{Q163*} and PRRT2^{Q250*} variants overexpressed in SH-SY5Y cells with and without NMD pathway inhibition.

* $p < 0.05$, *** $p < 0.001$ two-way ANOVA comparing PRRT2^{WT} and its variants, with *post hoc* Tukey multiple comparisons of means test. Data obtained from three technical and two biological replicates.

Comparison of the increase of mRNA levels between different variants and the wild-type after NMDI14 treatment showed a significant increase for PRRT2^{Q106*} ($p = 0.004$; Figure 4.23). The increase of mRNA levels for PRRT2^{Q106*} was approximately 5.5 fold, 3.4 fold for PRRT2^{Q163*} and 2.9 fold for PRRT2^{Q250*}, while the PRRT2^{WT} mRNA levels stayed almost unchanged.

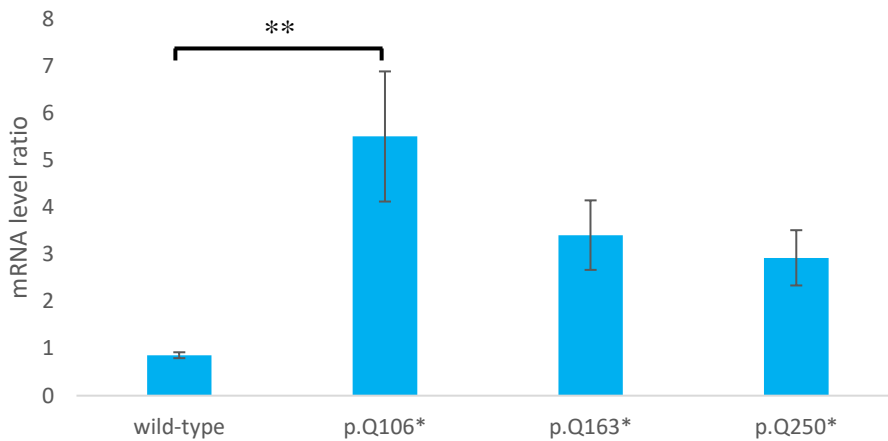


Figure 4.23. Ratio of mRNA levels change after treatment with NMDI14 for PRRT2^{WT} and PRRT2^{Q106*}, PRRT2^{Q163*} and PRRT2^{Q250*} variants overexpressed in SH-SY5Y cells.

** $p < 0.01$ one-way ANOVA comparing PRRT2^{WT} and its variants, with *post hoc* Tukey multiple comparisons of means test. Data obtained from three technical and two biological replicates.

The mean mRNA level ratio increase \pm SEM was 0.859 ± 0.063 for PRRT2^{WT}, 5.501 ± 1.381 for PRRT2^{Q106*}, 3.407 ± 0.737 for PRRT2^{Q163*} and 2.925 ± 0.587 for PRRT2^{Q250*}.

4.3.7. Role of variant PRRT2^{Q106*} in protein expression

Analysis of protein overexpression using the plasmids carrying the PRRT2^{WT} gene and the PRRT2^{Q106*}, PRRT2^{Q163*} and PRRT2^{Q250*} variants was performed using WB method and SH-SY5Y cells. No expression of PRRT2^{Q106*} was observed. Variants PRRT2^{Q163*} and PRRT2^{Q250*} both expressed significantly lower levels of protein than PRRT2^{WT} (PRRT2^{Q106*}, $p < 0.001$; PRRT2^{Q163*}, $p < 0.001$; PRRT2^{Q250*}, $p = 0.001$; Figure 4.24). The normalised mean band intensity \pm SEM was 1.000 ± 0.064 for PRRT2^{WT}, 0.0 ± 0.0 for PRRT2^{Q106*}, 0.121 ± 0.060 for PRRT2^{Q163*} and 0.311 ± 0.114 for PRRT2^{Q250*}.

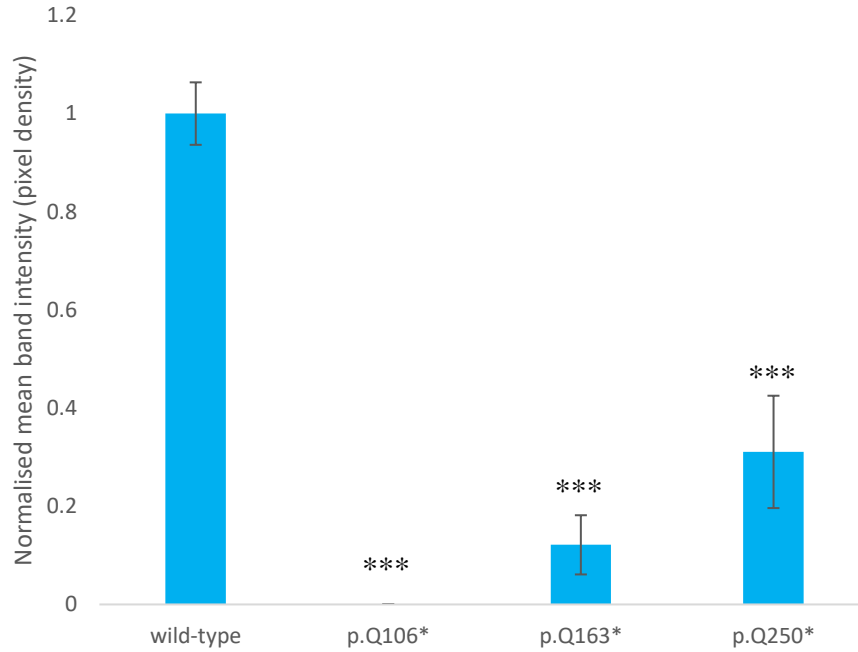


Figure 4.24. Western blot results for PRRT2^{WT} and PRRT2^{Q106*}, PRRT2^{Q163*} and PRRT2^{Q250*} variants overexpression in SH-SY5Y cells.

*** $p < 0.001$, one-way ANOVA comparing PRRT2^{WT} and its variants, with *post hoc* Tukey multiple comparisons of means test. Data obtained from three technical and biological replicates.

NMD pathway was inhibited by NMDI14 in the experimental group to test whether that would increase the truncated protein expression. No expression of PRRT2^{Q106*} was observed after NMD pathway inhibition, however, protein levels of the variants PRRT2^{Q163*} and PRRT2^{Q250*} increased significantly (from 0.121 ± 0.060 to 0.591 ± 0.120 , $p = 0.016$ and from 0.311 ± 0.114 to 1.127 ± 0.014 , $p < 0.001$ respectively). The protein levels of PRRT2^{WT} increased insignificantly (from 1.000 ± 0.064 to 1.371 ± 0.040 ; Figure 4.25).

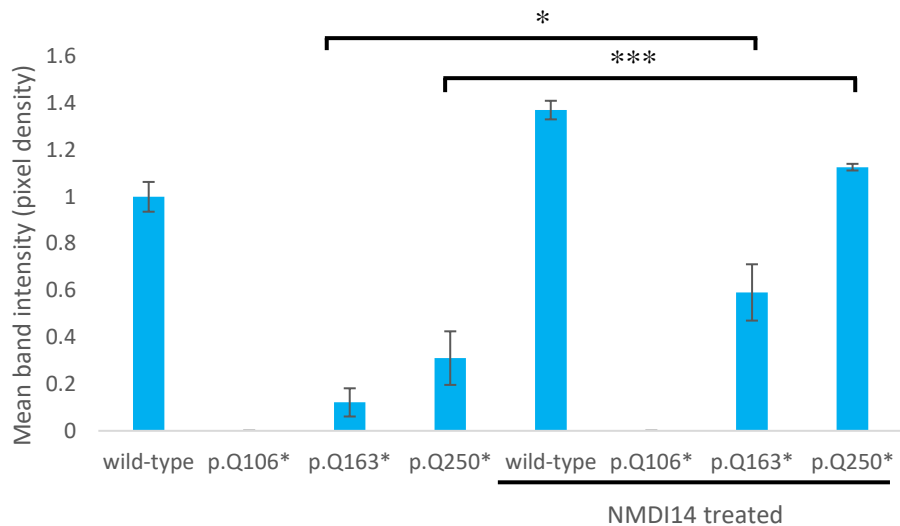


Figure 4.25. Western blot results for PRRT2^{WT} and PRRT2^{Q106*}, PRRT2^{Q163*} and PRRT2^{Q250*} variants overexpression in SH-SY5Y cells.

* $p < 0.05$, *** $p < 0.001$, two-way ANOVA comparing PRRT2^{WT} and its variants, with *post hoc* Tukey multiple comparisons of means test. Data obtained from three technical and biological replicates.

Comparison of the increase of protein levels between different variants and the wild-type after NMDI14 treatment showed no significant increase difference between the samples, although the increase of PRRT2^{Q163*} protein levels was approximately 5.8 fold and 4.2 fold for PRRT2^{Q250*}, while the PRRT2^{WT} protein levels stayed almost unchanged (Figure 4.26). The standard deviation was 0.180 for the PRRT2^{WT}, but as high as 2.676 and 2.260 for PRRT2^{Q163*} and PRRT2^{Q250*} respectively.

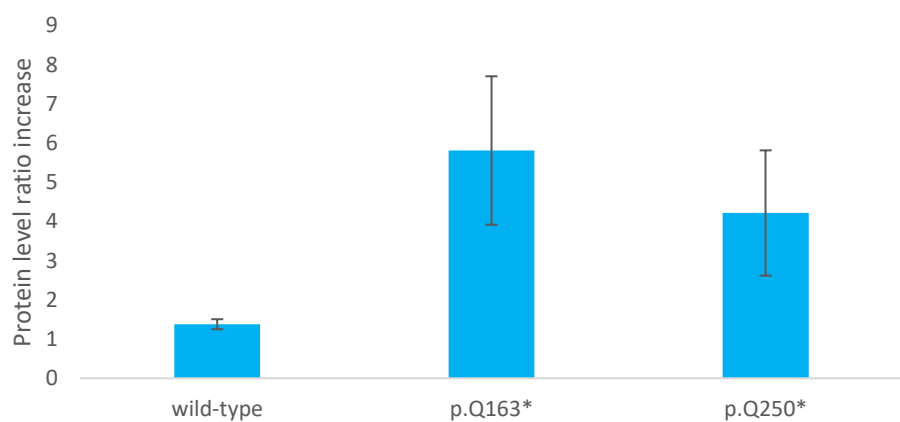


Figure 4.26. Ratio of protein levels increase after treatment with NMDI14 for PRRT2^{WT} and PRRT2^{Q163*} and PRRT2^{Q250*} variants overexpressed in SH-SY5Y cells.

5. DISCUSSION

5.1. Overview

The study of molecular mechanisms responsible for neurological diseases is important for development of possible treatments. Comprehending the complex nature of these diseases could allow us to focus research efforts on specific aspects of their molecular biology and find pathways for viable targeted treatment. AD, the most common neurodegenerative dementia, has been recognized as one of the most common causes of death in the United States by the late 70s (Katzman, 1976) and persists in the top ten causes to this day (Heron, 2019). PLS, although not as grave as ALS, forms a part of its pathological spectrum and significantly impairs the lives of affected patients. ALS is the most common progressive MND caused by neurodegeneration. PKD is a rare neurological disorder causing involuntary movements which negatively impacts the lives of affected patients. Although these diseases show such diverse symptoms, at the molecular level the mechanisms that cause them may sometimes be due to a SNP which affects haploinsufficiency, stability or function of the encoded protein.

Further, I will discuss the variants found which could possibly affect the diseases in three Spanish families affected by AD, PLS and PKD respectively. Considering the impact these variants may have on the protein they encode and what molecular mechanisms may be involved in causing these diseases. This should shed light on better comprehending the etiopathogenesis of these three neurological diseases.

5.2. Etiopathogenesis of AD in family UGM037

As three of the four siblings in family UGM037 were diagnosed with AD at ages 68, 69 and 72, it can be classified as late onset familial AD (Bird, 2018) which is typically associated with various susceptibility genes (Van Cauwenberghe *et al.*, 2015). The annual incidence rate (per 100 person-years) for AD is approximately 0.5 % for people aged 65 to 69 years and about 1 % for those aged 70 to 74 years (Mayeux and Stern, 2012). Therefore, although the patients are above the arbitrary line of 65 years, classifying them as having LOAD, they are at the lower range of that age classification. Among the siblings, one of the two men and both the women were diagnosed with AD, which is in accordance with data for differences between sexes (Lobo *et al.*, 2011). Although sex may be an AD risk

factor related to the fact that women, on average, live longer than men, here all the siblings lived to a similar age (between 85 and 89 years).

The next generation, children of the above mentioned siblings, were not diagnosed with AD, their DNA samples were collected at ages between 50 and 57. They may be at risk of developing the disease at a later age. Here, I will discuss how the AD susceptibility genes and the variant described in this work may contribute to the development of the disease in family UGM037.

5.2.1. Identification of a variant possibly responsible for AD in a Spanish family

WES was used to identify the variants possibly responsible for causing AD in UGM037 family. The exome contains approximately 85 % of known variants related to diseases, comprising only about 1 % of the genome (van Dijk *et al.*, 2014). This approach is a cost-effective method for identification of disease-related variants (Lacey *et al.*, 2014).

Variants within 24 AD susceptibility genes were identified through WES data analysis (Table 4.5), yet most of them were intronic and synonymous variants. Among them, only rs764542666 in gene *CR1* encoding a PTC (c.C406T, p.R136*) could be considered a damaging variant with a HIGH putative impact according to SnpEff. Analysis of all the variants available from WES data also pointed to the heterozygous rs764542666 SNV as damaging and most likely involved in causing the disease in the family. An introduction of a PTC early on (position 136) in a 2039 amino acid sequence suggests a significant impact on the function of the protein or its amount, affecting haploinsufficiency. Activation of the complement system is known to correspond with clinical expression of AD and the components of the system are found in amyloid plaques (Shen *et al.*, 2001). The complement receptor 1 was found to be involved in A β clearance (Zhu *et al.*, 2015) and a possible risk factor for AD through GWAS (Lambert *et al.*, 2009). Mahmoudi *et al.* (2015) suggested that it may be the lower levels of expression of the CR1*2 isoform which are responsible for higher AD risk, as this isoform was associated with AD susceptibility. This would align with the haploinsufficiency due to rs764542666 variant in *CR1* hypothesis.

Likely pathogenicity variant classification, according to Richards *et al.* (2015), and relatively lowest observed/expected score, according to Karczewski *et al.* (2020; <https://gnomad.broadinstitute.org/>), of the CR1^{R136*} variant strongly suggest its possible involvement in the disease and thus selecting it for further study. The segregation of the

variant with the disease and its rare occurrence in the general population, as well as in the samples from AD patients, is also hinting at its involvement in the disease development in family UGM037.

The role of variant CR1^{R136*} in pathophysiology of AD in family UGM037 is still undetermined. Lacking functional information, this variant could be in linkage disequilibrium with another variant that could be the real culprit.

As Lambert *et al.* (2009) showed, the SNV rs3818361 increases the risk for LOAD. The same effect could be observed for the CR1*2 isoform (Mahmoudi *et al.*, 2015). However, neither could be related to the disease in the family. Either they are absent, CR1*2, or do not segregate with the disease in the pedigree, rs3818361. Conversely a high incidence of APOE ϵ 4 was found in the family with all the affected individuals having at least one copy of ϵ 4 allele. A higher dose of APOE ϵ 4 is related to decreasing the age of onset and is a major risk factor in families with LOAD (Corder *et al.*, 1993). The same study suggests that having a homozygous APOE ϵ 4/ ϵ 4 genotype is “virtually sufficient to cause AD by age 80”. This statement was refuted by Henderson *et al.* (1995), showing the existence of cognitively normal ϵ 4/ ϵ 4 carriers at an older age. However, APOE ϵ 4/ ϵ 4 genotype remains a strong risk factor for developing the disease. In family UGM037 two out of three affected individuals were homozygous for this allele with the third affected individual being heterozygous. Although none of the patients’ onsets can be qualified as early (< 65 years), the ones homozygous for APOE ϵ 4 are reported to have been diagnosed at an earlier age (68, 69 years old) than the heterozygous one (72 years old). According to Mayeux and Stern (2012) the annual incidence rate (per 100 person-years) of AD increases from 0.30 % at ages 65-69 to 0.85 % at ages 70-74. The mean age of onset for individuals with APOE ϵ 4/ ϵ 4 is 68 ± 8.2 years, for individuals with APOE ϵ 3/ ϵ 4 76 ± 8.2 years and for individuals with APOE ϵ 3/ ϵ 3 84 ± 8.2 years (Hane *et al.*, 2018). Therefore, the difference between the ages of onset of the patients is probably insignificant. Unexpectedly, the unaffected individual who did not carry the CR1^{R136*} variant, had no signs of cognitive decline and carried the ϵ 4/ ϵ 4 genotype at the APOE locus, passing at 88 years of age. This supports Henderson *et al.* (1995) claim, refuting the assertion by Corder *et al.* (1993), that homozygous APOE ϵ 4/ ϵ 4 genotype is enough to cause AD by 80 years of age. Lack of CR1^{R136*} variant in this individual may imply it is this SNV that causes the disease in this family and the APOE ϵ 4 dose only increases the chances of developing AD and contributes to earlier onset.

While autosomal dominant mutations in genes *APP*, *PSEN1* and *PSEN2* mostly contribute to EO-fAD (Cruts *et al.*, 2012), the SNV CR1^{R136*} seems to affect the patients causing LO-fAD. I propose the heterozygous PTC nature of this SNV causes haploinsufficiency and a similar effect as the lower levels of expression of CR1*2 isoform which is said to be connected to increased complement activation and lower A β clearance (Hazrati *et al.*, 2012; Mahmoudi *et al.*, 2015).

Among the younger family members, whose DNA samples were collected at ages 50 to 57 years, three of seven had the CR1^{R136*} variant. Among the three, none of them had the rs3818361 variant nor the CR1*2 isoform. Two of them had *APOE* ϵ 3/ ϵ 4 genotype and one had *APOE* ϵ 3/ ϵ 3 genotype. Reiteration of these results is important as these members of the family could face a risk of developing AD at an older age. The presence of one copy of the *APOE* ϵ 4 allele increases the risk of AD 4-fold (Tanzi, 2012), which applies to six of the four younger family members, and two also carry the rs3818361 variant. As the contribution of variant CR1^{R136*} in pathophysiology of AD in family UGM037 is still not clear, it can only be speculated whether it will negatively affect the younger family members. However, in line with the clinical and WES data for the older family members, the patients with CR1^{R136*} variant and *APOE* ϵ 3/ ϵ 4 genotype should be at the highest risk of developing the disease. If the variant in CR1 is itself causative of the disease and the *APOE* ϵ 4 dose lowers the onset age, the individual with *APOE* ϵ 3/ ϵ 3 genotype could also develop AD, but possibly at an older age. Although these are only speculations, they may be beneficial to both the individuals and clinicians as early detection and treatment can enhance quality of life (Small, 2000).

The role variant CR1^{R136*} plays in mRNA and protein expression is important to understand the possible molecular mechanisms by which it may be involved in causing AD. In an *in vitro* cellular model, the mRNA and protein levels of CR1^{R136*} variant were significantly lower than those of CR1^{WT}. NMD pathway is often responsible for destruction of mRNA with a PTC. The early position of the PTC in CR1 suggests the involvement of the NMD pathway in destruction of the transcript (Coban-Akdemir *et al.*, 2018). The data from experimental NMD inhibition support the hypothesis that it is through this pathway that CR1^{R136*} transcript is decayed, not allowing for significant production of CR1 protein. This corroborates the theory of loss-of-function of CR1 and further haploinsufficiency possibly being involved in AD. Therefore, I suggest the transcript is decayed in NMD pathway and not enough fully functional protein can be expressed to control the complement system and help with A β clearance. The A β increase provokes activation of

the complement system which has been reported by Singharo *et al.* (2000) to be harmful to neurons. The CR1*2 isoform is expressed at lower protein levels than CR1*1, and it is assumed that it may be the lower levels of the protein that affect the risk of AD (Hazarti *et al.*, 2012; Mahmoudi *et al.*, 2015). The hypothesis of haploinsufficiency would thus be in line with the previous assumptions.

The *in vitro* overexpression of a protein in a cellular model has its limitations. Among the general are differences between the cellular model and the corresponding cells in the organism, problems with establishing appropriate microenvironment, such as interactions with other cells, or the fact that the protein is artificially overexpressed. Due to the lack of data on interaction of CR1 with its ligands C3b nor C4b, the scope of loss-of-function cannot be fully established. Further research is needed to confirm no gain-of-function is caused by the variant in CR1. Immunocytochemistry study could provide information on whether the truncated protein is expressed and accumulated in the cells. Also, fibroblasts collected from patients and converted into neuronal cells or cellular model altered through gene editing could be used to more closely represent the natural metabolism of the cell. Further functional studies on the effects of CR1^{R136*} variant should be done, as its segregation with the disease in the pedigree may be coincidental and the main culprit for development of AD in family UGM037 may be age and incidence of *APOE* ϵ 4. The variant CR1^{R136*} is very rare and not well studied with regards to its effect on AD, therefore, it would be interesting to further investigate this potentially causative mutation and its role in developing the disease.

5.3. Etiopathogenesis of PLS in family UGM471

The three members of family UGM471 were diagnosed with PLS at ages of 24, 28 and 30 years. This rare disease, with incidence of 1 in 10 million per year (Brugman and Wokke, 2004), while having an earlier average age of onset than ALS, has a mean onset age of 54.6 ± 10.9 years (Tartaglia *et al.*, 2007). The onset of the symptoms in the affected family members is therefore early, and apparently can be classified as limb-onset. Limb-onset is more common than bulbar among PLS than ALS patients, and in general it is responsible for 70 % of cases among ALS patients (Zarei *et al.*, 2015). Therefore, the feeling of stiffness spreading from one leg to the second and later upper extremities in the patients is in line with the limb-onset and the typical progression of the disease.

5.3.1. Identification of a variant possibly responsible for PLS in a Spanish family

Two variants were identified which could possibly be the cause of the symptoms in the family. One was a novel variant c.G884C p.R295P in *ADPRH* and the other a previously described mutation c.T497C p.L166P in *PSENI* (rs63750265). The latter was discovered within the UGM471 family genome recently Vazquez-Costa (forthcoming), therefore they will be discussed chronologically, from discovery.

5.3.1.1. Variant *ADPRH*^{R295P} in Spanish family affected by primary lateral sclerosis

Variants within 18 ALS and PLS susceptibility genes were identified through WES data analysis (Table 4.13), yet most of them were intron and synonymous variants. No variants within these genes were considered possibly deleterious and most were previously described and assigned an accession number in the dbSNP, having high frequency within the general population according to gnomAD (Karczewski *et al.*, 2020; <https://gnomad.broadinstitute.org/>). Interestingly no common variants were found in genes *SOD1* nor *TARDBP* in all of the affected individuals. The more specifically related to PLS rather than ALS genes, *SPG7* and *TBKI* (Yang *et al.*, 2016; Gomez-Tortosa *et al.*, 2017) did not yield interesting results either. About 76 % of familial cases of ALS can be attributed to a known ALS-related gene (Morgan and Orell, 2016), whereas there is no such data on PLS. However, the WES results did not show a clear candidate variant possibly responsible for affecting the disease in the family.

A general assessment of the WES data was done to identify possible variants which could be involved in the disease and variant c.G884C p.R295P in *ADPRH* was revealed as the most interesting through filtering, prioritization and segregation. Its deleterious/damaging impact according to SIFT and Polyphen predictions and its involvement in many essential cellular pathways are implications for a further investigation. The heterogeneous pathology of ALS and thus PLS, which involves many mechanisms in the cell, suggests there may be more undiscovered risk factors and mutations that could affect the disease. The ADP-ribosylation cycle involving *ADPRH* has important functions in cellular regulation. It is involved in proliferation, differentiation, DNA repair, and other essential cellular processes (Palazzo *et al.*, 2019). *ADPRH* is involved in tumorigenesis and bacterial response to toxins (Mashimo *et al.*, 2014), however, mutations in *ADPRS*, which shares 41 % similarity with *ADPRH*, are known to

be associated with neurodegeneration (Danhauser *et al.*, 2018; Mashimo *et al.*, 2019). It is important to point out ADPRH is involved in arginine-specific MARYlation while ADPRS in serine-specific PARYlation (Rack *et al.*, 2018; Rack *et al.*, 2020).

Sanger sequencing validation of the WES results confirmed the segregation of variant ADPRH^{R295P} with the disease in the pedigree (Figure 4.9). In addition, the same variant was present in two unaffected members of this family. This particular variant was not previously described although other amino acid changes at the same position are present in the dbSNP (c.G884A p.R295Q (rs150984649; 26/282,798 allele count/number), c.G884T p.R295L (rs150984649; 1/31,398) and c.C883T p.R295X (rs190409478; 149/282,764)). All of them were found to be rare, at frequencies < 0.0006. Although the frequency of ADPRH^{R295P} variant could only be studied in a limited control population of 196 samples, it was not present in any of them. This, together with rare frequency of variants at position R295 and ADPRH^{R295P} being a novel variant suggests it is also a rare SNV.

The observed/expected score for gene constraint reported by gnomAD (Karczewski *et al.*, 2020; <https://gnomad.broadinstitute.org/>) for missense variants in *ADPRH* suggest it is relatively tolerant for such changes. However, effects of proline mutations can be quite severe for protein stability. Proline does not possess an important amide hydrogen and can cause structural rigidity of the protein structure due to the ϕ dihedral angle being constrained to values around -65° . This, depending on the location of proline change, can significantly destabilize or cause aggregation of a protein (Bajaj *et al.*, 2007). Here, the proline replaces a conserved arginine in an α -helix, which is predicted to disturb its structure and influence the function and activity of the protein. Arginine at position 295 creates salt bridges with aspartic acid at position 280 and glutamic acid at position 291, which are predicted to be interrupted due to a change in charge when replaced with a proline. Stability analysis also predicts this change to be very destabilizing, with more than twice as much free energy change than for other described missense variants in the same position.

A variant in a paralog of *ADPRH*, c.A881C p.L294R *ADPRHL1* (rs41306688) was identified to segregate with the disease in the family, not being present in any of the healthy individuals, not even the ones with ADPRH^{R295P} variant. In *Xenopus*, ADPRH and ADPRHL1 may share actin as their target, working in unison to allow for correct actin filament assembly (Smith *et al.*, 2020). Giampetruzzi *et al.* (2019) suggest modulation of actin polymerisation may be involved in disruption of the nucleocytoplasmic transport

which is important in ALS pathogenesis. It affects nuclear import, mRNA post-transcriptional regulation and the stability and function of the nuclear pore. The variant ADPRHL1^{L294R} is predicted to be damaging/deleterious according to Polyphen and SIFT, although it is not as rare as the novel variant in ADPRH with a frequency of around 0.02, according to gnomAD (5619/250,028). Although ADPRHL1 is mainly restricted to the heart (Smith *et al.*, 2016), its RNA expression is ubiquitous (Stelzer *et al.*, 2016; www.genecards.org). Its catalytic activity as an ADP-ribose hydrolase is said to be lost (Oka *et al.*, 2006). However, as mentioned before, the newest data suggest it may have a role in actin filament assembly in *Xenopus* (Smith *et al.*, 2020). No interaction between ADPRH and ADPRHL1 could be confirmed through co-IP, however, this approach has many limitations, as low-affinity protein interactions may not be detected, the antibody chosen may affect the outcome of the assay and the experimental conditions do not represent the intricate inner workings of the cell. Also, ADPRH and ADPRHL1 may work together without directly interacting with one another.

The study of ADPRH^{R295P} activity, revealed it to be just as efficient as ADPRH^{WT} at cleaving ADP-ribose from agmatine. The reaction of ADP-ribosylation of agmatine using NAD⁺ as an ADP-ribose donor, promoted by CT, causes a significant increase in ADP-ribosyl-agmatine in comparison with spontaneous ADP-ribosylation of agmatine. Using UPLC for separation and identification of components the retention time for ADP-ribosyl-agmatine was assigned at 1.40 minutes. This allowed for correct interpretation and quantification of the ADPRH activity assay results. The ADPRH activity efficiency was reduced in the double variant ADPRH^{D55A/D56A} where two point mutations, changed two of the essential aspartic acid residues to alanines. This result refutes the *in silico* prediction of disruption of proper activity of the protein. However, this may be due to the conditions of the reaction used. The protein was overexpressed in *E. coli DH5α* bacterial strain for extraction and purification and not in eukaryotic cells. Protein expression was induced by IPTG and chaperonin GroEL/GroES was co-expressed to counter protein misfolding and promote bacterial survival. The reaction itself was performed *in vitro* and not *in vivo*, which may have influenced the accuracy of the result. Therefore, this assay does not take into account the prediction of a destabilizing effect of the variant on the protein. If it is in fact as strong as predicted, a hypothesis of haploinsufficiency could be put forward.

While mRNA expression of ADPRH^{R295P} proved to be as high as that of ADPRH^{WT} and variants ADPRH^{R295Q} and ADPRH^{D55A/D56A}, the protein levels of the novel SNV were significantly lower, suggesting an important destabilizing effect. As expected, the stability

of mRNA in the missense variants was not affected. However, the protein stability seems to be altered greatly by ADPRH^{R295P}. Interestingly, other missense variants, whether at the same (ADPRH^{R295Q}) or at a different position (ADPRH^{D55A/D56A}), did not cause a significant loss of protein. The *in silico* analysis of the novel variant ADPRH^{R295P} suggests the arginine at this position is well conserved and a change to a proline could affect the protein stability which seems to be consistent with our results. The protein levels of variant ADPRH^{R295Q} at the same position are not significantly different from ADPRH^{WT} although they show consistently lower band signal on a WB. At the same time the double variant ADPRH^{D55A/D56A} seems to have very similar protein levels as ADPRH^{WT}. This may suggest a change at position 295 is important for protein stability. The ADPRH^{R295*} SNV showed lower mRNA levels and no protein could be detected with WB. This may suggest that the NMD pathway was involved in decay of ADPRH^{R295*} mRNA, since it encodes a truncated protein with a PTC. The truncating variant caused a complete lack of ADPRH which may be attributed to its low mRNA levels, as well as further misfolding or instability of the protein if it was to be translated.

Although this work could confirm a strong destabilizing effect of the novel variant in *ADPRH*, it is difficult to quantify how much of the protein is lost due to that effect. Further study would need to be performed to determine whether it could contribute to haploinsufficiency, preferably using fibroblasts collected from patients and converted into neuronal cells or a cellular model altered through gene editing. Also further studies on the relation between *ADPRH*, *ADPRHL1*, their variants and actin could answer whether they could have an impact on causing PLS. The fact that healthy members of the family were carriers of ADPRH^{R295P} could be a case of lower penetrance of this variant or the need for both variants in *ADPRH* and *ADPRHL1* in order for the pathological phenotype to manifest. Although it is difficult to say if the novel variant ADPRH^{R295P} affects the disease in the Spanish family, little is known about MARylation, its effects on diseases and molecular mechanisms it is involved in, therefore it may be important to gather this information for improvement of fundamental research.

5.3.1.2. Variant PSEN1^{L166P} in Spanish family affected by primary lateral sclerosis

Recently Jazi *et al.* (2019) have associated mutations PSEN1^{A431E} (rs63750083) and PSEN1^{L381V} (rs63750687) with PLS and Vazquez-Costa (forthcoming) has identified a

mutation c.T497C p.L166P in *PSEN1* with the accession number rs63750265 in the PLS patients from family UGM471.

Although mutations in *PSEN1* are commonly causing EOAD (Tanzi, 2012) there have been reports associating *PSEN1* with ALS (Panas *et al.*, 2000; Couthouis *et al.*, 2014). *PSEN1*^{A431E} is an AD causing mutation associated with a founder effect in the Mexican state of Jalisco (Yescas *et al.*, 2006). The two patients reported by Jazi *et al.* (2019) with this mutation and diagnosed with PLS were from a neighbouring state of Michoacán. They developed FTD, spastic quadriparesis, bulbar dysfunction and spastic anarthria at 40 – 44 years of age. The age of onset for the individuals affected by this mutation was between 33 and 44 years. Although cognitive problems were the usual symptoms, there were also motor symptoms recorded, such as myoclonus, spastic paraplegia and pyramidal rigidity (Yescas *et al.*, 2006). One family recorded with *PSEN1*^{A431E} and AD, had mostly motor symptoms with atypical cognitive changes (Santos-Mandujano *et al.*, 2019). *PSEN1*^{L381V} is another mutation related to AD which Jazi *et al.* (2019) recorded in a PLS patient. Once again it is related with an early onset at ages 29 - 32, and motor symptoms, spastic paraparesis (Mehrabian *et al.*, 2004; Ikeuchi *et al.*, 2008; Dintchov Traykov *et al.*, 2009). Couthouis *et al.* (2014) have associated variant *PSEN1*^{W203C} (rs1384308168) with ALS through a targeted sequencing screen, however, not much is known about the biological influence this variant could exert.

The mutant *PSEN1*^{L166P} (rs63750265), found in the Spanish family UGM471, was confirmed to segregate with the disease in the pedigree, not being present in any of the healthy individuals. Other mutations at the same position were described and associated with AD. *PSEN1*^{L166V} was discovered in a British person with typical AD symptoms onset at age 42 and death 8 years later (Sassi *et al.*, 2014). *PSEN1*^{L166R} (rs63750265) was described in a family of Spanish origin (Ezquerria *et al.*, 2000). The age of onset was 32 - 44 years and some of the symptoms included cognitive decline, grasping reflex, bradykinesia, aphasia and akinetic movements. *PSEN1*^{L166H} (rs63750265) was described in an Italian woman with age of onset at 30 years (Pantieri *et al.*, 2005). Her symptoms included cognitive decline, myoclonus and postural tremor. *PSEN1*^{L166del} (rs63751458) was described in a woman with age of onset at 38 years (Knight *et al.*, 2007). Her symptoms were mostly related to cognitive impairment. Mutant *PSEN1*^{L166P} (rs63750265) is a rare mutation associated with spastic paraparesis and AD with a very early age of onset. It was first described in a 15 year old woman with generalised seizures. As the disease progressed she showed symptoms of depression, memory loss, ataxia and spastic paraparesis with

moderate dementia by the age 28 and death at 35. The post-mortem examination showed she had A β plaques in her cerebral cortex (Moehlmann *et al.*, 2002). Another carrier of PSEN1^{L166P} mutation was a person who, at age 23, showed symptoms of memory impairment, spastic paraparesis and dysarthria. Here, the post-mortem examination showed A β and tau pathology (Lyoo *et al.*, 2016).

The PSEN1^{L166P} mutation causes partial loss of γ -secretase cleavage function and increases the A β ₄₂/A β ₄₀ ratio by reducing the A β ₄₀ levels (Koch *et al.*, 2012). This is a result of decreasing the cleavage of A β ₄₃ to A β ₄₀. While A β ₄₂ levels are not affected, the disruption of the γ -secretase cleavage in the line A β ₄₉-A β ₄₀, which digests A β ₄₉ into shorter peptides, induces higher levels of longer A β peptides and reduction of the shorter ones. This in turn increases the A β ₄₂/A β ₄₀ ratio (Li *et al.*, 2016). An *in vitro* experiment by Chávez-Gutiérrez *et al.* (2012) estimated that PSEN1^{L166P} mutation decreased the cleavage efficiency of γ -secretase by 75 %. This may be due to increased dissociation of γ -secretase, with the mutation PSEN1^{L166P} present, from A β peptides not allowing for cleavage (Szaruga *et al.*, 2017). PSEN1^{L166P} was found to disrupt the proteins conformation (Berezovska *et al.*, 2005). It is also associated with an increase in APP β -C-terminal fragments (β -CTFs) which are a biomarker for AD in the CSF (García-Ayllón *et al.*, 2017; Kwart *et al.*, 2019). Interestingly, PSEN1^{L166P} was also shown to have a negative effect on PSEN1^{WT}. An *in vitro* study showed PSEN1^{WT} A β production was suppressed in the presence of PSEN1^{L166P} possibly due to hetero-oligomerization (Zhou *et al.*, 2017). In addition, presenilins form endoplasmic reticulum (ER) Ca²⁺ leak channels, accounting for approximately 80 % of passive Ca²⁺ leak. This function was proven to be impaired in several fAD-linked mutations, with PSEN1^{L166P} among them (Tu *et al.*, 2006; Nelson *et al.*, 2007).

Recent discovery of the PSEN1^{L166P} mutation in the UGM471 family did not allow for a more thorough investigation of its effects, but it was confirmed to segregate with the disease in the pedigree. The aggressive nature of this mutation, the early age of onset and the motor symptoms, strongly suggest it is PSEN1^{L166P} which causes the disease in the family. While it is not associated with ALS nor PLS, the patients may have been also affected by other environmental or genetic factors to produce this phenotype. At the same time Jazi *et al.* (2019) report PLS patients with mutations in *PSEN1*. Either the effect of the *PSEN1* mutations is heterogeneous enough to cause these different disorders, the disorders are much more related due to the molecular mechanisms that cause them, or the effect of *PSEN1* mutation is affected by interaction with other proteins. In a familial setting,

environmental or other genetic factors may also contribute to PSEN1^{L166P} causing a PLS phenotype. The destabilising effect of ADPRH^{R295P} together with the predicted damaging effect of ADPRHL1^{L294R} may be the genetic contributor to this phenotypic variability. Further investigation into molecular mechanisms could shed light onto this subject. Certainly, common neurodegeneration-linked genes should be looked at when identifying a possible cause of a disease, regardless of which disease they are associated with. Patients with PLS, ALS or spastic paraparesis should be investigated for *PSEN1* mutations.

5.4. Etiopathogenesis of PKD in family UGM478

Patients diagnosed with PKD in family UGM478 had the age of onset between 8 and 14 years which is in line with data reported by Schelosky (2010). Four out of five were male and they are either in remission or without attack altogether at an older age, with or without the carbamazepine treatment. The sex ratio aligns with the 3-4:1 described in other studies (Palau-Bargues *et al.*, 2010; Singer *et al.*, 2010; Roze *et al.*, 2015), although McGuire *et al.* (2018) suggest no such disparity exists in the familial form of the disease. They also communicate the typical trend of the symptoms decreasing in severity or even resolving completely in adulthood, which is in line with what can be observed in patients from the UGM478 family. Sudden movement was identified as the most common trigger among these patients which indicates the patients having the kinesigenic type of PxD (Demirkiran and Jankovic, 1995).

5.4.1. Identification of a variant possibly responsible for PKD in a Spanish family

Variants within 7 genes associated with PxDs were identified through WES data analysis (Table 4.30) and another 8 variants from all the WES data. Most of them were intronic and synonymous variants. Among the genes associated with PxDs, *PRRT2* which is the only gene whose mutations are a known cause for PKD (Chen *et al.*, 2011), was identified with 5 different variants. One was a synonymous variant, two were upstream gene variants, one was a 3'-UTR variant and one was a stop gain variant. All except for the stop gain variant were previously reported in dbSNP. Truncating mutations in *PRRT2* are a common cause of PKD (Méneret *et al.*, 2013; Valtorta *et al.*, 2016) and the same novel stop gain variant c.C316T p.Q106* in *PRRT2* was identified by filtering and prioritization of the entire WES data. This novel variant was considered damaging with a HIGH putative

impact according to SnpEff. PRRT2^{Q106*} segregated with the disease in the pedigree and could be classified as pathogenic according to Richards *et al.* (2015). The observed/expected score for gene constraint reported by gnomAD (Karczewski *et al.*, 2020; <https://gnomad.broadinstitute.org/>) for nonsense variants in *PRRT2* suggest it is intolerant for such changes, which is in accordance with the truncating mutations in this gene being a common cause of PKD (Méneret *et al.*, 2013; Valtorta *et al.*, 2016). PRRT2^{Q106*} is a novel variant and could not be found in any database. The ASPCR frequency estimation showed it was absent in our small control population and while it segregated with the disease in the pedigree, it was not present in any of the healthy family members. The penetrance of *PRRT2* mutations in PKD is estimated at 61 – 65 % (Marini *et al.*, 2012; Van Vliet *et al.*, 2012), however, taking into account the infantile convulsions phenotype, the penetrance is close to 100 % (Van Vliet *et al.*, 2012). PKD, BIFE and ICCA are not only caused by the same mutations in *PRRT2* but can concurrently exist in different members of the same family (Yang *et al.*, 2013). These are all strong indicators that PRRT2^{Q106*} is a novel mutation causing PKD in Spanish family UGM478.

It is suggested that the mechanism by which the PTCs affect PKD is by *PRRT2* haploinsufficiency due to NMD (Gardiner *et al.*, 2012; Wu *et al.*, 2014). In accordance with their results, significantly lower mRNA levels for the variants with PTCs (PRRT^{Q106*}, PRRT^{Q163*}, PRRT^{Q250*}), in comparison with PRRT^{WT}, were found in a cellular model with *PRRT2* and its variants overexpressed (Figure 4.21). This confirmed the previous results for variant PRRT^{Q163*}, studied by Wu *et al.* (2014). Results for PRRT^{Q106*} and PRRT^{Q250*} are also in line with that study. Inhibition of the NMD pathway by treatment of transfected SH-SY5Y cells with NMDI14, increased the mRNA levels for PRRT^{Q106*} and PRRT^{Q163*} and showed an increasing trend for PRRT^{Q250*}. This suggests that the NMD pathway may in fact be the culprit behind mRNA decay prompted by PTCs in the variants studied. Interestingly protein levels of the novel variant PRRT^{Q106*} were undetectable with WB before and after NMDI14 treatment, while the other variants studied (PRRT^{Q163*} and PRRT^{Q250*}) had significantly lower protein levels than PRRT^{WT} before the treatment and increased levels after. As outlined by Williamson (1994), the proline-rich regions are often important for protein stability. I propose the reason behind the lack of protein encoded by the PRRT^{Q106*} was due to a combined effect of NMD pathway destruction of mRNA and destabilization of the protein truncated ahead of the stabilizing proline-rich region. These results suggest the novel variant PRRT^{Q106*} is probably the cause of PKD in the UGM478 Spanish family. The molecular mechanisms responsible for

the affliction may be the NMD pathway causing decay of the transcript leading to haploinsufficiency. Lack of PRRT2 in turn causes hyperexcitability through dysregulated neurotransmitter release and hyperactivity of Na⁺ channels.

The general limitations of the *in vitro* overexpression of a protein in a cellular model apply here. Differences between the cellular model and the corresponding cells in the organism, problems with establishing appropriate microenvironment, such as interactions with other cells, or the fact that the protein is artificially overexpressed in naturally unavailable amounts. Due to the lack of data on interaction between PRRT2 with SNAP-25, GluA1 or Na_v1.2 and Na_v1.6 channels, scope of loss-of-function cannot be fully established. Similarly as with the recommendations for the truncating variant in *CR1* in AD family (section 5.2) further research is needed to confirm no gain-of-function is caused by the variant in PRRT2. Immunocytochemistry study could provide information on whether the truncated protein is expressed and accumulated in the cells. Fibroblasts collected from patients and converted into neuronal cells or cellular model altered through gene editing could be used to more closely represent the natural metabolism of the cell. Also, functional studies on the effects of PRRT2^{Q106*} variant could be done. Mutations in *PRRT2* are well established causes of PKD, and these results support the claim that it is the NMD pathway leading to haploinsufficiency that causes the disease in family UGM478. Further investigation would be recommended to comprehend the full scope of PRRT2 function and therefore the molecular mechanisms which are involved in causing the disease when haploinsufficiency is in action.

5.5. Final remarks

Common or related pathological molecular mechanisms may affect neurological disorders, traditionally considered as unrelated, in the intricate network of the nervous system. In this work I have outlined some of such putative mechanisms. Single nucleotide variants which may affect different phenotypes through partial loss-of-function due to protein destabilisation or haploinsufficiency due to NMD.

Dang *et al.* (2008) have identified 299 human haploinsufficient genes by database-mining and text-searching in PubMed and OMIM over 10 years ago. They found 6 % of them to be associated with neurological disorders. Only two years later Huang *et al.* (2010) have estimated the possible number of human haploinsufficient genes to be 12,443 out of approximately 22,000. While the total number of human genes is a matter of debate and

further study (Salzberg, 2018), their estimation indicates that we can expect a great number of genes where protein level dose effect may be essential. Many studies have shown how important haploinsufficiency is in neurological disorders (Kobayashi *et al.*, 2003; Freischmidt *et al.*, 2015; Cheng *et al.*, 2019). Recently *C9orf72* was found to be haploinsufficient in ALS/FTD due to the GGGGCC repeat expansion (Shi *et al.*, 2018).

In this work, I am postulating that *CR1* and *PRRT2* are haploinsufficient in Spanish families with AD and PKD respectively. While it is established that most mutations in *PRRT2* lead to loss-of-function and haploinsufficiency (Fruscione *et al.*, 2018), to my knowledge there is no such reports on *CR1*. Haploinsufficiency, therefore, emerges as a common factor between these and other neurological diseases. Furthermore, the molecular mechanism behind the *CR1* and *PRRT2*-related haploinsufficiency seems to be the NMD elicited by SNVs encoding PTCs. Single base-pair substitutions account for approximately 67 % of human disease-associated mutations (Antonarakis and Cooper, 2013). Nonsense mutations make up ~ 20 % of those substitutions, and ~ 11 % of all human disease-causing mutations (Mort *et al.*, 2008). NMD is predicted to be elicited by 49 % of nonsense SNPs while the other 51 % would lead to expression of a truncated protein (Yamaguchi-Kabata *et al.* 2008). The novel variant in *PRRT2* and variant rs764542666 in *CR1* appear to contribute to the former percentage, demonstrating a common molecular mechanism in distinct neurological diseases.

Loss-of-function is strictly related to haploinsufficiency which is a dominant phenotype in organisms heterozygous for such alleles (Deutschbauer *et al.*, 2005). Although variant PSEN1^{L166P} (rs63750265) was not found to cause haploinsufficiency, it affects a partial loss of γ -secretase cleavage (Koch *et al.*, 2012) and ER Ca²⁺ leak channel function (Nelson *et al.*, 2007). The variant ADPRH^{R295P}, identified in the same family, is shown to significantly destabilize the protein, drastically affecting its levels. This in turn may impede its function. Wang and Moulton (2001) show that 80 % of disease-associated missense mutations affect protein stability. Whether ADPRH^{R295P} variant in heterozygosis is in fact deleterious remains to be seen, depending on its tolerance to decreased protein dose. However, loss-of-function, whether full or partial, encompasses the underlying molecular mechanisms of the SNVs described in this work, which contribute to independent neurological diseases.

Discussing the molecular mechanisms in AD and PLS in the two Spanish families, and the involvement of CR1^{R136*} (rs764542666) and PSEN1^{L166P} (rs63750265) variants in disease pathogenesis, it is important not to omit other possibly contributing factors. While

CR1^{R136*} may be a causative mutation in the family, its members had a high incidence of *APOE* ϵ 4. As mentioned, a higher dose of *APOE* ϵ 4 is related to decreasing the age of onset and is a major risk factor in families with LOAD (Corder *et al.*, 1993). Hardy (1994) found that in families with *APP* mutations, the incidence of *APOE* ϵ 4 was related to an earlier age of onset, while the incidence of *APOE* ϵ 2, with a later age of onset, with regards to *APOE* ϵ 3. Interestingly, Benitez *et al.* (2013) showed PSEN1^{E318G} variant to be related to an increased risk of AD, dependent on *APOE* ϵ 4. While otherwise PSEN1^{E318G} was considered non-pathogenic, its interaction with *APOE* ϵ 4 increased A β deposition, causing a faster cognitive decline and neurodegeneration. Thus, while carrying CR1^{R136*} variant may be sufficient to develop AD, it is also probable that the members of the Spanish family studied, were affected solely by the *APOE* ϵ 4 risk factor or a combination of the two.

Similarly, other factors, whether environmental or genetic, may affect symptoms developed by the family with PLS. Although PSEN1^{L166P} seems to be responsible for the phenotype experienced by the patients, their symptoms differ from the more canonical AD features related to this variant (Moehlmann *et al.*, 2002; Lyoo *et al.*, 2016). This divergent disease expression may be due to *PSEN1* gene pleiotropy (Ibanez *et al.*, 2018), however, it may also be due to other contributing factors. Environmental factors have been found to play an important role in ALS (Oskarsson *et al.*, 2015) and they cannot be disregarded in a familial disease. Here, I propose a genetic factor which may contribute to the dissimilar symptoms experienced by the members of this family. Novel variant ADPRH^{R295P} may have an effect on the disease development, destabilising actin filaments in the presence of ADPRHL1^{L294R} variant, prompting a phenotype closer to PLS, together with the aggressive PSEN1^{L166P} mutation. The complexity of neurological diseases comes, in part, from a cumulative nature of defects that cause them, and thus it is always essential to search for other factor which may add to the observed phenotype. Further studies into the effects ADPRH variants may have on neurological diseases are needed as it contributes to the still poorly understood, but very important MARylation.

6. CONCLUSIONS

1. Variant rs764542666 encoding a PTC (c.C406T, p.R136*) is likely the first known Alzheimer's disease causative mutation in *CRI*, based on WES and genetic expression study. NMD pathway provoked haploinsufficiency is its probable molecular mechanism, causing lower rate of A β clearance and higher complement system activation. It may be sufficient to cause LOAD or contributes to its development in *APOE* ϵ 4 allele presence, encouraging research into the rare truncating variants in this gene.

2. Mutant rs63750265 in *PSENI* encoding a missense mutation (c.T497C, p.L166P) is the likely cause of primary lateral sclerosis, based on WES study, segregation analysis and previous knowledge, raising questions of pleiotropic effects of the mutation in a gene typically causing AD. Loss of γ -secretase cleavage function, increase of A β ₄₂/A β ₄₀ ratio and impairment of ER Ca²⁺ leak channel function are the probable molecular mechanisms behind the mutation. Strongly destabilising variant in *ADPRH* (c.G884C, p.R295P) may be a factor contributing to the observed phenotypical variability.

3. Novel variant in gene *PRRT2* encoding a PTC (c.C316T, p.Q106*) is the likely cause of paroxysmal kinesigenic dyskinesia, based on WES and genetic expression study. This work supports the hypothesis of NMD pathway provoking haploinsufficiency of *PRRT2* as the molecular mechanism behind PKD, leading to hyperexcitability and hyperactivity of Na⁺ channels.

4. Single nucleotide variants in different genes may lead to partial or full loss-of-function due to effects on protein conformation, stability or by NMD causing haploinsufficiency. Together with other contributing factor, these are common molecular mechanisms which are crucial in study of complex neurological diseases.

7. REFERENCES

- Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., *et al.* (2010). A method and server for predicting damaging missense mutations. *Nature Methods*, 7(4), 248–249. doi:10.1038/nmeth0410-248
- Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., *et al.* (2011). The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, 7(3), 270–279. Doi:10.1016/j.jalz.2011.03.008
- Alcaz, S., Jarebinski, M., Pekmezović, T., Marinković, Z., & Apostolski, S. (1997). Amiotroficka lateralna skleroza u definisanoj populaciji Beograda–epidemioloska studija [Amyotrophic lateral sclerosis in a defined population in Belgrade—an epidemiologic study]. *Vojnosanit Pregl.* 54(1), 19-26.
- Alonso, A. D., Cohen, L. S., Corbo, C., Morozova, V., ElIdrissi, A., Phillips, G., & Kleiman, F. E. (2018). Hyperphosphorylation of Tau Associates With Changes in Its Function Beyond Microtubule Stability. *Frontiers in Cellular Neuroscience*, 12. Doi:10.3389/fncel.2018.00338
- Alzheimer, A. (1907). Uber eine eigenartige Erkrankung der Hirnrinde. *Allgemeine Zeitschrift fur Psychiatrie und phychish-Gerichtliche Medizin, (Berlin)* 64: 146-148.
- Alzheimer's Association. (2018). 2018 Alzheimer's disease facts and figures. *Alzheimer's Dement.* 14, 367–429.
- Antonarakis, S. E., & Cooper, D. N. (2013). Human Gene Mutation in Inherited Disease. Emery and Rimoin's Principles and Practice of Medical Genetics, 1–48. doi:10.1016/b978-0-12-383834-6.00007-0
- Arthur, K. C., Calvo, A., Price, T. R., Geiger, J. T., Chiò, A., & Traynor, B. J. (2016). Projected increase in amyotrophic lateral sclerosis from 2015 to 2040. *Nature Communications*, 7, 12408. Doi:10.1038/ncomms12408
- Ashkenazy, H., Abadi, S., Martz, E., Chay, O., Mayrose, I., Pupko, T., & Ben-Tal, N. (2016). ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Research*, 44(W1), W344–W350. Doi:10.1093/nar/gkw408

- Auton, A., Abecasis, G., Altshuler, D. *et al.* (2015). A global reference for human genetic variation. *Nature*, 526, 68–74. <https://doi.org/10.1038/nature15393>
- Awada, A. (2015). Early and late-onset Alzheimer's disease: What are the differences? *Journal of Neurosciences in Rural Practice*, 6(3), 455. Doi:10.4103/0976-3147.154581
- Bajaj, K., Madhusudhan, M. S., Adkar, B. V., Chakrabarti, P., Ramakrishnan, C., Sali, A., & Varadarajan, R. (2007). Stereochemical Criteria for Prediction of the Effects of Proline Mutations on Protein Stability. *PLoS Computational Biology*, 3(12), e241. doi:10.1371/journal.pcbi.0030241
- Ballard, C., Gauthier, S., Corbett, A., Brayne, C., Aarsland, D., & Jones, E. (2011). Alzheimer's disease. *The Lancet*, 377(9770), 1019–1031. Doi:10.1016/s0140-6736(10)61349-9
- Barber, S. C., & Shaw, P. J. (2010). Oxidative stress in ALS: Key role in motor neuron injury and therapeutic target. *Free Radical Biology and Medicine*, 48(5), 629–641. Doi:10.1016/j.freeradbiomed.2009.11.018
- Barnum, S. R. (1995). Complement Biosynthesis in the Central Nervous System. *Critical Reviews in Oral Biology & Medicine*, 6(2), 132–146. Doi:10.1177/10454411950060020301
- Becker, J. T., Boller, F., Lopez, O. L., Saxton, J., & McGonigle, K. L. (1994). The Natural History of Alzheimer's Disease. *Archives of Neurology*, 51(6), 585. Doi:10.1001/archneur.1994.00540180063015
- Beek, J., Elward, K., & Gasque, P. (2003). Activation of Complement in the Central Nervous System. *Annals of the New York Academy of Sciences*, 992(1), 56–71. Doi:10.1111/j.1749-6632.2003.tb03138.x
- Benitez, B. A., Karch, C. M., Cai, Y., Jin, S. C., Cooper, B., *et al.* (2013). The PSEN1, p.E318G Variant Increases the Risk of Alzheimer's Disease in APOE-ε4 Carriers. *PLoS Genetics*, 9(8), e1003685. doi:10.1371/journal.pgen.1003685
- Berezovska, O. (2005). Familial Alzheimer's Disease Presenilin 1 Mutations Cause Alterations in the Conformation of Presenilin and Interactions with Amyloid Precursor Protein. *Journal of Neuroscience*, 25(11), 3009–3017. doi:10.1523/jneurosci.0364-05.2005
- Berman, H. M. (2000). The Protein Data Bank. *Nucleic Acids Research*, 28(1), 235–242. Doi:10.1093/nar/28.1.235

- Bermejo-Pareja, F., Benito-León, J., Vega, S., Medrano, M. J., & Román, G. C. (2008). Incidence and subtypes of dementia in three elderly populations of central Spain. *Journal of the Neurological Sciences*, 264(1-2), 63–72. Doi:10.1016/j.jns.2007.07.021
- Berthold, C. L., Wang, H., Nordlund, S., & Hogbom, M. (2009). Mechanism of ADP-ribosylation removal revealed by the structure and ligand complexes of the dimanganese mono-ADP-ribosylhydrolase DraG. *Proceedings of the National Academy of Sciences*, 106(34), 14247–14252. Doi:10.1073/pnas.0905906106
- Bertram, L., & Tanzi, R. E. (2005). The genetic epidemiology of neurodegenerative disease. *Journal of Clinical Investigation*, 115(6), 1449–1457. Doi:10.1172/jci24761
- Bertram, L., Lange, C., Mullin, K., Parkinson, M., Hsiao, M., Hogan, M. F., *et al.* (2008). Genome-wide Association Analysis Reveals Putative Alzheimer's Disease Susceptibility Loci in Addition to APOE. *The American Journal of Human Genetics*, 83(5), 623–632. Doi:10.1016/j.ajhg.2008.10.008
- Biedler, J. L., Roffler-Tarlov, S., Schachner, M., & Freedman, L. S. (1978). Multiple neurotransmitter synthesis by human neuroblastoma cell lines and clones. *Cancer Res.* 38(11 Pt 1), 3751-3757.
- Bin, Y., Zhu, Q., Li, M., & Xia, J. (2019). Comprehensive Analysis of Alzheimer's Disease Biologically Candidate Causal Genes Revealed by Function Association Study with GWAS. *IEEE Access*, 1–1. Doi:10.1109/access.2019.2935515
- Binder, L. I. (1985). The distribution of tau in the mammalian central nervous system. *The Journal of Cell Biology*, 101(4), 1371–1378. Doi:10.1083/jcb.101.4.1371
- Bird TD. Alzheimer Disease Overview. 1998 Oct 23 [Updated 2018 Dec 20]. In: Adam MP, Ardinger HH, Pagon RA, *et al.*, editors. GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle; 1993-2020. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK1161/>
- Bobinski, M., Wegiel, J., Wisniewski, H. M., Tarnawski, M., Bobinski, M., Reisberg, B., *et al.* (1996). Neurofibrillary pathology — correlation with hippocampal formation atrophy in Alzheimer disease. *Neurobiology of Aging*, 17(6), 909–919. Doi:10.1016/s0197-4580(97)85095-6
- Bonafede, R., & Mariotti, R. (2017). ALS Pathogenesis and Therapeutic Approaches: The Role of Mesenchymal Stem Cells and Extracellular Vesicles. *Frontiers in*

- Cellular Neuroscience*, 11. Doi:10.3389/fncel.2017.00080
- Braak, H., & Braak, E. (1991). Neuropathological staging of Alzheimer-related changes. *Acta Neuropathologica*, 82(4), 239–259. Doi:10.1007/bf00308809
- Braschi, B., Denny, P., Gray, K., Jones, T., Seal, R., Tweedie, S., *et al.* (2018). Genenames.org: the HGNC and VGNC resources in 2019. *Nucleic Acids Research*. Doi:10.1093/nar/gky930
- Brugman, F., Wokke, J. H. J. (2004). Primary lateral sclerosis. *Orphanet Encyclopedia*, <https://www.orpha.net/data/patho/GB/uk-PLS.pdf>
- Brugman, F., Eymard-Pierre, E., van den Berg, L. H., Wokke, J. H. J., Gauthier-Barichard, F., & Boespflug-Tanguy, O. (2007). Adult-onset primary lateral sclerosis is not associated with mutations in the *ALS2* gene *Neurology*, 69(7), 702–704. Doi:10.1212/01.wnl.0000267322.48364.76
- Bruyn, R. P., Koelman, J. H., Troost, D., & de Jong, J. M. (1995). Motor neuron disease (amyotrophic lateral sclerosis) arising from longstanding primary lateral sclerosis. *Journal of Neurology, Neurosurgery & Psychiatry*, 58(6), 742–744. Doi:10.1136/jnnp.58.6.742
- Caille, I., Allinquant, B., Dupont, E., Bouillot, C., Langer, A., Müller, U., & Prochiantz, A. (2004). Soluble form of amyloid precursor protein regulates proliferation of progenitors in the adult subventricular zone. *Development*, 131(9), 2173–2181. Doi:10.1242/dev.01103
- Camacho, A., Esteban, J., & Paradas, C. (2018). Report by the Spanish Foundation for the Brain on the social impact of amyotrophic lateral sclerosis and other neuromuscular disorders. *Neurología (English Edition)*, 33(1), 35–46. Doi:10.1016/j.nrleng.2015.02.006
- Capriotti, E., Fariselli, P., & Casadio, R. (2005). I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. *Nucleic Acids Research*, 33(Web Server), W306–W310. Doi:10.1093/nar/gki375
- Cellini, E., Tedde, A., Bagnoli, S., Pradella, S., Piacentini, S., Sorbi, S., & Nacmias, B. (2009). Implication of Sex and *SORL1* Variants in Italian Patients With Alzheimer Disease. *Archives of Neurology*, 66(10). Doi:10.1001/archneurol.2009.101
- Charcot, J. M., & Joffroy, A. (1869). Deux cas d’atrophie musculaire progressive avec lésions de la substance grise et des faisceaux antéro-latéraux de la moelle

- épinière. *Arch Physiol Neurol Path.* 2, 744
- Chasseigneaux, S., & Allinquant, B. (2011). Functions of A β , sAPP α and sAPP β : similarities and differences. *Journal of Neurochemistry*, 120, 99–108. Doi:10.1111/j.1471-4159.2011.07584.x
- Chávez-Gutiérrez, L., Bammens, L., Benilova, I., Vandersteen, A., Benurwar, M., Borgers, M., *et al.* (2012). The mechanism of γ -Secretase dysfunction in familial Alzheimer disease. *The EMBO Journal*, 31(10), 2261–2274. doi:10.1038/emboj.2012.79
- Chen, S., Sayana, P., Zhang, X., & Le, W. (2013). Genetics of amyotrophic lateral sclerosis: an update. *Molecular Neurodegeneration*, 8(1), 28. Doi:10.1186/1750-1326-8-28
- Chen, W.-J., Lin, Y., Xiong, Z.-Q., Wei, W., Ni, W., Tan, G.-H., *et al.* (2011). Exome sequencing identifies truncating mutations in PRRT2 that cause paroxysmal kinesigenic dyskinesia. *Nature Genetics*, 43(12), 1252–1255. Doi:10.1038/ng.1008
- Cheng, A., Wang, J., Ghena, N., Zhao, Q., Perone, I., King, M. T., *et al.* (2019). SIRT3 Haploinsufficiency Aggravates Loss of GABAergic Interneurons and Neuronal Network Hyperexcitability in an Alzheimer's Disease Model. *The Journal of Neuroscience*, 1446–19. doi:10.1523/jneurosci.1446-19.2019
- Chiò, A., Logroscino, G., Traynor, B. J., Collins, J., Simeone, J. C., Goldstein, L. A., & White, L. A. (2013). Global Epidemiology of Amyotrophic Lateral Sclerosis: A Systematic Review of the Published Literature. *Neuroepidemiology*, 41(2), 118–130. Doi:10.1159/000351153
- Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Sayers, E. W. (2015). GenBank. *Nucleic Acids Research*, 44(D1), D67–D72. Doi:10.1093/nar/gkv1276
- Clement, A. M., Nguyen, M. D., Roberts, E. A., Garcia, M. L., Boillée, S., Rule, M., McMahon, A. P., Doucette, W., Siwek, D., Ferrante, R. J., Brown Jr, R. H., Julien, J.-P., Goldstein, L. S. B., & Cleveland, D. W. (2003). Wild-Type Nonneuronal Cells Extend Survival of SOD1 Mutant Motor Neurons in ALS Mice. *Science*, 302(5642), 113–117. Doi:10.1126/science.1086071
- Coban-Akdemir, Z., White, J. J., Song, X., Jhangiani, S. N., Fatih, J. M., Gambin, T., *et al.* (2018). Identifying Genes Whose Mutant Transcripts Cause Dominant

- Disease Traits by Potential Gain-of-Function Alleles. *The American Journal of Human Genetics*, 103(2), 171–187. Doi:10.1016/j.ajhg.2018.06.009
- Coleman, J., Jouannot, O., Ramakrishnan, S. K., Zanetti, M. N., Wang, J., Salpietro, V., *et al.* (2018). PRRT2 Regulates Synaptic Fusion by Directly Modulating SNARE Complex Assembly. *Cell Reports*, 22(3), 820–831. Doi:10.1016/j.celrep.2017.12.056
- Corcia, P., Couratier, P., Blasco, H., Andres, C. R., Beltran, S., Meininger, V., & Vourc'h, P. (2017). Genetics of amyotrophic lateral sclerosis. *Revue Neurologique*, 173(5), 254–262. Doi:10.1016/j.neurol.2017.03.030
- Corder, E., Saunders, A., Strittmatter, W., Schmechel, D., Gaskell, P., Small, G., *et al.* (1993). Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science*, 261(5123), 921–923. Doi:10.1126/science.8346443
- Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Research*, 16(22), 10881–10890. Doi:10.1093/nar/16.22.10881
- Coupé, C., & Gordon, P. H. (2013). Amyotrophic Lateral Sclerosis – Clinical Features, Pathophysiology and Management. *European Neurological Review*, 8(1), 38–44. Doi: <http://doi.org/10.17925/ENR.2013.08.01.38>
- Couthouis, J., Raphael, A. R., Daneshjou, R., & Gitler, A. D. (2014). Targeted Exon Capture and Sequencing in Sporadic Amyotrophic Lateral Sclerosis. *PLoS Genetics*, 10(10), e1004704. doi:10.1371/journal.pgen.1004704
- Cruts, M., Theuns, J., & Van Broeckhoven, C. (2012). Locus-specific mutation databases for neurodegenerative brain diseases. *Human Mutation*, 33(9), 1340–1344. Doi:10.1002/humu.22117
- Cukier, H. N., Kunkle, B. K., Hamilton, K. L., *et al.* (2017). Exome Sequencing of Extended Families with Alzheimer's Disease Identifies Novel Genes Implicated in Cell Immunity and Neuronal Function. *Journal of Alzheimer's Disease & Parkinsonism*, 07(04). Doi:10.4172/2161-0460.1000355
- Culbertson, M. R. (1999). RNA surveillance: unforeseen consequences for gene expression, inherited genetic disorders and cancer. *Trends in Genetics*, 15(2), 74–80. Doi:10.1016/s0168-9525(98)01658-8
- Da Cruz, S., & Cleveland, D. W. (2011). Understanding the role of TDP-43 and FUS/TLS in ALS and beyond. *Current Opinion in Neurobiology*, 21(6), 904–

919. Doi:10.1016/j.conb.2011.05.029

- Dang, V. T., Kassahn, K. S., Marcos, A. E., & Ragan, M. A. (2008). Identification of human haploinsufficient genes and their genomic proximity to segmental duplications. *European Journal of Human Genetics*, 16(11), 1350–1357. doi:10.1038/ejhg.2008.111
- Danhauser, K., Alhaddad, B., Makowski, C., Piekutowska-Abramczuk, D., Syrbe, S., Gomez-Ospina, N., *et al.* (2018). Bi-allelic ADPRHL2 Mutations Cause Neurodegeneration with Developmental Delay, Ataxia, and Axonal Neuropathy. *The American Journal of Human Genetics*. Doi:10.1016/j.ajhg.2018.10.005
- Dawson, N. L., Lewis, T. E., Das, S., Lees, J. G., Lee, D., Ashford, P., *et al.* (2016). CATH: an expanded resource to predict protein function through structure and sequence. *Nucleic Acids Research*, 45(D1), D289–D295. Doi:10.1093/nar/gkw1098
- De Strooper, B. (2003). Aph-1, Pen-2, and Nicastrin with Presenilin Generate an Active γ -Secretase Complex. *Neuron*, 38(1), 9–12. Doi:10.1016/s0896-6273(03)00205-8
- De Vos, K. J., Chapman, A. L., Tennant, M. E., Manser, C., Tudor, E. L., Lau, K.-F., *et al.* (2007). Familial amyotrophic lateral sclerosis-linked SOD1 mutants perturb fast axonal transport to reduce axonal mitochondria content. *Human Molecular Genetics*, 16(22), 2720–2728. Doi:10.1093/hmg/ddm226
- De Vries, B. S., Rustemeijer, L. M. M., van der Kooi, A. J., Raaphorst, J., Schröder, C. D., Nijboer, T. C. W., *et al.* (2017). A case series of PLS patients with frontotemporal dementia and overview of the literature. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, 18(7-8), 534–548. Doi:10.1080/21678421.2017.1354996
- Demirkiran, M., & Jankovic, J. (1995). Paroxysmal dyskinesias: Clinical features and classification. *Annals of Neurology*, 38(4), 571–579. Doi:10.1002/ana.410380405
- Deutschbauer, A. M., Jaramillo, D. F., Proctor, M., Kumm, J., Hillenmeyer, M. E., Davis, R. W., Nislow, C., & Giaever, G. (2005). Mechanisms of haploinsufficiency revealed by genome-wide profiling in yeast. *Genetics*, 169(4), 1915–1925. <https://doi.org/10.1534/genetics.104.036871>

- Dintchov Traykov, L., Mehrabian, S., Van den Broeck, M., Radoslavova Raycheva, M., Cruts, M., Kirilova Jordanova, A., & Van Broeckhoven, C. (2009). Novel PSEN1 Mutation in a Bulgarian Patient With Very Early-Onset Alzheimer's Disease, Spastic Paraparesis, and Extrapyrimal Signs. *American Journal of Alzheimer's Disease & Other Dementiasr*, 24(5), 404–407. doi:10.1177/1533317509341464
- Dodou, E., Xu, S. M., & Black, B. L. (2003). Mef2c is activated directly by myogenic basic helix-loop-helix proteins during skeletal muscle development in vivo. *Mech. Dev.* 120,1021 -1032.
- Dölle, C., Rack, J. G. M., & Ziegler, M. (2013). NAD and ADP-ribose metabolism in mitochondria. *FEBS Journal*, 280(15), 3530–3541. Doi:10.1111/febs.12304
- Dorszewska, J., Prendecki, M., Oczkowska, A., Dezor, M., & Kozubski, W. (2016). Molecular Basis of Familial and Sporadic Alzheimer's Disease. *Current Alzheimer Research*, 13(9). Doi: 10.2174/1567205013666160314150501
- Drown, B. S., Shirai, T., Rack, J. G. M., Ahel, I., & Hergenrother, P. J. (2018). Monitoring Poly(ADP-ribosyl)glycohydrolase Activity with a Continuous Fluorescent Substrate. *Cell Chemical Biology*. Doi:10.1016/j.chembiol.2018.09.008
- Dubois, B., Padovani, A., Scheltens, P., Rossi, A., & Dell'Agnello, G. (2015). Timely Diagnosis for Alzheimer's Disease: A Literature Review on Benefits and Challenges. *Journal of Alzheimer's Disease*, 49(3), 617–631. Doi:10.3233/jad-150692
- Durazzo, T. C., Mattsson, N., & Weiner, M. W. (2014). Smoking and increased Alzheimer's disease risk: A review of potential mechanisms. *Alzheimer's & Dementia*, 10(3), S122–S145. Doi:10.1016/j.jalz.2014.04.009
- Dyle, M. C., Kolakada, D., Cortazar, M. A., & Jagannathan, S. (2019). How to get away with nonsense: mechanisms and consequences of escape from nonsense-mediated RNA decay. *Wiley Interdiscip. Rev. RNA*. 11(1), e1560. <https://doi.org/10.1002/wrna.1560>
- Ebrahimi-Fakhari, D., Moufawad El Achkar, C., & Klein, C. (2018). PRRT2-Associated Paroxysmal Movement Disorders. In: Adam, M. P., Ardinger, H. H., Pagon, R. A., *et al.*, editors. GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle; 1993-2020. Available from:

<https://www.ncbi.nlm.nih.gov/sites/books/NBK475803/>

- Eikelenboom, P., & Stam, F. C. (1982). Immunoglobulins and complement factors in senile plaques. *Acta Neuropathologica*, 57(2-3), 239–242. Doi:10.1007/bf00685397
- Endo, Y., Matsushita, M., & Fujita, T. (2011). The role of ficolins in the lectin pathway of innate immunity. *The International Journal of Biochemistry & Cell Biology*, 43(5), 705–712. Doi:10.1016/j.biocel.2011.02.003
- Erro, R., & Bhatia, K. P. (2018). Unravelling of the paroxysmal dyskinesias. *Journal of Neurology, Neurosurgery & Psychiatry*, jnnp–2018–318932. Doi:10.1136/jnnp-2018-318932
- Evans, D. A., Funkenstein, H. H., Albert, M. S., *et al.* (1989). Prevalence of Alzheimer's Disease in a Community Population of Older Persons. *JAMA*, 262(18), 2551. Doi:10.1001/jama.1989.03430180093036
- Ezquerro, M., Carnero, C., Blesa, R., & Oliva, R. (2000). A Novel Presenilin 1 Mutation (Leu166Arg) Associated With Early-Onset Alzheimer Disease. *Archives of Neurology*, 57(4), 485. Doi:10.1001/archneur.57.4.485
- Ferri, C. P., Prince, M., Brayne, C., Brodaty, H., Fratiglioni, L., Ganguli, M., *et al.* (2005). Global prevalence of dementia: a Delphi consensus study. *The Lancet*, 366(9503), 2112–2117. Doi:10.1016/s0140-6736(05)67889-0
- Finegan, E., Chipika, R. H., Shing, S. L. H., Hardiman, O., & Bede, P. (2019). Primary lateral sclerosis: a distinct entity or part of the ALS spectrum? *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, 1–13. Doi:10.1080/21678421.2018.1550518
- Flirski, M., Sieruta, M., Golanska, E., Kłoszewska, I., Liberski, P. P., & Sobow, T. M. (2012). PRND 3'UTR polymorphism may be associated with 193behavioural disturbances in Alzheimer disease. *Prion*, 6(1), 73–80. Doi:10.4161/pri.6.1.18428
- Freischmidt, A., Wieland, T., Richter, B., Ruf, W., Schaeffer, V., Müller, K., *et al.* (2015). Haploinsufficiency of TBK1 causes familial ALS and fronto-temporal dementia. *Nature Neuroscience*, 18(5), 631–636. doi:10.1038/nn.4000
- Fruscione, F., Valente, P., Sterlini, B., Romei, A., Baldassari, S., Fadda, M., *et al.* (2018). PRRT2 controls neuronal excitability by negatively modulating Na⁺ channel 1.2/1.6 activity. *Brain*, 141(4), 1000–1016. Doi:10.1093/brain/awy051

- García-Ayllón, M.-S., Lopez-Font, I., Boix, C. P., Fortea, J., Sánchez-Valle, R., Lleó, A., *et al.* (2017). C-terminal fragments of the amyloid precursor protein in cerebrospinal fluid as potential biomarkers for Alzheimer disease. *Scientific Reports*, 7(1). doi:10.1038/s41598-017-02841-7
- Gardiner, A. R., Bhatia, K. P., Stamelou, M., Dale, R. C., Kurian, M. A., Schneider, S. A., *et al.* (2012). PRRT2 gene mutations: From paroxysmal dyskinesia to episodic ataxia and hemiplegic migraine. *Neurology*, 79(21), 2115–2121. doi:10.1212/wnl.0b013e3182752c5a
- Gasque, P., Dean, Y. D., McGreal, E. P., VanBeek, J., & Morgan, B. P. (2000). Complement components of the innate immune system in health and disease in the CNS. *Immunopharmacology*, 49(1-2), 171–186. Doi:10.1016/s0162-3109(00)80302-1
- Ghani, M., Reitz, C., George-Hyslop, P. S., & Rogaeva, E. (2018). Genetic Complexity of Early-Onset Alzheimer's Disease. *Neurodegenerative Diseases*, 29–50. Doi:10.1007/978-3-319-72938-1_3
- Ghosh, S. G., Becker, K., Huang, H., Salazar, T. D., Chai, G., Salpietro, V., *et al.* (2018). Biallelic Mutations in ADPRHL2, Encoding ADP-Ribosylhydrolase 3, Lead to a Degenerative Pediatric Stress-Induced Epileptic Ataxia Syndrome. *The American Journal of Human Genetics*. Doi:10.1016/j.ajhg.2018.07.010
- Giampetruzzi, A., Danielson, E. W., Gumina, V., Jeon, M., Boopathy, S., Brown, R. H., *et al.* (2019). Modulation of actin polymerization affects nucleocytoplasmic transport in multiple forms of amyotrophic lateral sclerosis. *Nature Communications*, 10(1). doi:10.1038/s41467-019-11837-y
- Gijselinck, I., Van Mossevelde, S., van der Zee, J., Sieben, A., Engelborghs, S., *et al.* (2015). The C9orf72 repeat size correlates with onset age of disease, DNA methylation and transcriptional downregulation of the promoter. *Molecular Psychiatry*, 21(8), 1112–1124. Doi:10.1038/mp.2015.159
- Gireud, M., Sirisaengtaksin, N., & Bean, A. J. (2014). Molecular Mechanisms of Neurological Disease. *From Molecules to Networks*, 639–661. Doi:10.1016/b978-0-12-397179-1.00021-x
- Giri, M., Lü, Y., & Zhang, M. (2016). Genes associated with Alzheimer's disease: an overview and current status. *Clinical Interventions in Aging*, 665. Doi:10.2147/cia.s105769

- Goetz, C. G. (2000). Amyotrophic lateral sclerosis: Early contributions of Jean-Martin Charcot. *Muscle & Nerve*, 23(3), 336–343. Doi:10.1002/(sici)1097-4598(200003)23:3<336::aid-mus4>3.0.co;2-l
- Goltermann, L., Sarusie, M. V., & Bentin, T. (2016). Chaperonin GroEL/GroES Over-Expression Promotes Aminoglycoside Resistance and Reduces Drug Susceptibilities in Escherichia coli Following Exposure to Sublethal Aminoglycoside Doses. *Frontiers in Microbiology*, 6. Doi:10.3389/fmicb.2015.01572
- Gómez-Pinedo, U., Galán, L., Matías-Guiu, J. A., Pytel, V., Moreno, T., Guerrero-Sola, A., & Matías-Guiu, J. (2019). Notch Signalling in the Hippocampus of Patients With Motor Neuron Disease. *Frontiers in Neuroscience*, 13. Doi:10.3389/fnins.2019.00302
- Gómez-Tortosa, E., Van der Zee, J., Ruggiero, M., Gijssels, I., Esteban-Pérez, J., García-Redondo, A., *et al.* (2017). Familial primary lateral sclerosis or dementia associated with Arg573GlyTBK1 mutation. *Journal of Neurology, Neurosurgery & Psychiatry*, 88(11), 996–997. Doi:10.1136/jnnp-2016-315250
- Gordon, P. H., Cheng, B., Katz, I. B., Pinto, M., Hays, A. P., Mitsumoto, H., & Rowland, L. P. (2006). The natural history of primary lateral sclerosis. *Neurology*, 66(5), 647–653. Doi:10.1212/01.wnl.0000200962.94777.71
- Gordon, P. H., Salachas, F., Lacomblez, L., Le Forestier, N., Pradat, P.-F., Bruneteau, G., *et al.* (2012). Predicting Survival of Patients with Amyotrophic Lateral Sclerosis at Presentation: A 15-Year Experience. *Neurodegenerative Diseases*, 12(2), 81–90. Doi:10.1159/000341316
- Gordon, P. (2013). Amyotrophic Lateral Sclerosis: An update for 2013 Clinical Features, Pathophysiology, Management and Therapeutic Trials. *Aging and Disease*, 04(05), 295–310. Doi:10.14336/ad.2013.0400295
- Grundke-Iqbal, I., Iqbal, K., Tung, Y. C., Quinlan, M., Wisniewski, H. M., & Binder, L. I. (1986). Abnormal phosphorylation of the microtubule-associated protein tau (tau) in Alzheimer cytoskeletal pathology. *Proceedings of the National Academy of Sciences*, 83(13), 4913–4917. Doi:10.1073/pnas.83.13.4913
- Grunnet, M., Jespersen, T., Angelo, K., Frøkjær-Jensen, C., Klaerke, D. A., Olesen, S.-P., & Jensen, B. S. (2001). Pharmacological modulation of SK3 channels. *Neuropharmacology*, 40(7), 879–887. Doi:10.1016/s0028-3908(01)00028-4

- Guerreiro, R., & Bras, J. (2015). The age factor in Alzheimer's disease. *Genome Medicine*, 7(1). Doi:10.1186/s13073-015-0232-5
- Han, M.-R., Schellenberg, G. D., & Wang, L.-S. (2010). Genome-wide association reveals genetic effects on human A β 42 and τ protein levels in cerebrospinal fluids: a case control study. *BMC Neurology*, 10(1). Doi:10.1186/1471-2377-10-90
- Hane, F., Augusta, C., & Bai, O. (2018). A hierarchical Bayesian model to predict APOE4 genotype and the age of Alzheimer's disease onset. *PLoS ONE* 13(7): e0200263. <https://doi.org/10.1371/journal.pone.0200263>
- Hardy, J., & Higgins, G. (1992). Alzheimer's disease: the amyloid cascade hypothesis. *Science*, 256(5054), 184–185. Doi:10.1126/science.1566067
- Hardy, J. (1994). Genetics of Alzheimer's disease: APP, APOE and the chromosome 14 locus. *Neuropsychopharmacology*, 11(4), 260–260. doi:10.1038/sj.npp.1380119
- Harold, D., Abraham, R., Hollingworth, P., Sims, R., Gerrish, A., Hamshere, M. L., *et al.* (2009). Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. *Nature Genetics*, 41(10), 1088–1093. Doi:10.1038/ng.440
- Hazrati, L.-N., Van Cauwenberghe, C., Brooks, P. L., Brouwers, N., Ghani, M., Sato, C., *et al.* (2012). Genetic association of CR1 with Alzheimer's disease: A tentative disease mechanism. *Neurobiology of Aging*, 33(12), 2949.e5–2949.e12. doi:10.1016/j.neurobiolaging.2012.07.001
- Hedera, P., Xiao, J., Puschmann, A., Momčilović, D., Wu, S. W., & LeDoux, M. S. (2012). Novel *PRRT2* mutation in an African-American family with paroxysmal kinesigenic dyskinesia. *BMC Neurology*, 12(1). Doi:10.1186/1471-2377-12-93
- Henderson, A., Jorm, A., Korten, A., Christensen, H., Jacomb, P., Easteal, S., *et al.* (1995). Apolipoprotein E allele ϵ 4, dementia, and cognitive decline in a population sample. *The Lancet*, 346(8987), 1387–1390. Doi:10.1016/s0140-6736(95)92405-1
- Heron M. (2019). Deaths: Leading causes for 2017. *National Vital Statistics Reports*, 68, 6. Hyattsville, MD: National Center for Health Statistics..
- Holers, V. M., Chaplin, D. D., Leykam, J. F., Gruner, B. A., Kumar, V., & Atkinson, J.

- P. (1987). Human complement C3b/C4b receptor (CR1) mRNA polymorphism that correlates with the CR1 allelic molecular weight polymorphism. *Proceedings of the National Academy of Sciences*, 84(8), 2459–2463. Doi:10.1073/pnas.84.8.2459
- Howland, D. S., Liu, J., She, Y., Goad, B., Maragakis, N. J., Kim, B., *et al.* (2002). Focal loss of the glutamate transporter EAAT2 in a transgenic rat model of SOD1 mutant-mediated amyotrophic lateral sclerosis (ALS). *Proceedings of the National Academy of Sciences*, 99(3), 1604–1609. Doi:10.1073/pnas.032539299
- Huang, N., Lee, I., Marcotte, E. M., & Hurles, M. E. (2010). Characterising and Predicting Haploinsufficiency in the Human Genome. *PLoS Genetics*, 6(10), e1001154. doi:10.1371/journal.pgen.1001154
- Huat, T. J., Camats-Perna, J., Newcombe, E. A., Valmas, N., Kitazawa, M., & Medeiros, R. (2019). Metal Toxicity Links to Alzheimer’s Disease and Neuroinflammation. *Journal of Molecular Biology*. Doi:10.1016/j.jmb.2019.01.018
- Hudson, A. J., Kiernan, J. A., Munoz, D. G., Pringle, C. E., Brown, W. F., & Ebers, G. C. (1993). Clinicopathological features of primary lateral sclerosis are different from amyotrophic lateral sclerosis. *Brain Research Bulletin*, 30(3-4), 359–364. Doi:10.1016/0361-9230(93)90265-d
- Hug, N., Longman, D., & Cáceres, J. F. (2016). Mechanism and regulation of the nonsense-mediated decay pathway. *Nucleic Acids Research*, 44(4), 1483–1495. Doi:10.1093/nar/gkw010
- Ibanez, L., Dube, U., Davis, A. A., Fernandez, M. V., Budde, J., Cooper, B., *et al.* (2018). Pleiotropic Effects of Variants in Dementia Genes in Parkinson Disease. *Frontiers in Neuroscience*, 12. doi:10.3389/fnins.2018.00230
- Iida, K., & Nussenzweig, V. (1981). Complement receptor is an inhibitor of the complement cascade. *Journal of Experimental Medicine*, 153(5), 1138–1150. Doi:10.1084/jem.153.5.1138
- Ikeuchi, T., Kaneko, H., Miyashita, A., Nozaki, H., Kasuga, K., Tsukie, T., *et al.* (2008). Mutational Analysis in Early-Onset Familial Dementia in the Japanese Population. *Dementia and Geriatric Cognitive Disorders*, 26(1), 43–49. doi:10.1159/000141483

- Jack, C. R., Albert, M. S., Knopman, D. S., McKhann, G. M., Sperling, R. A., Carrillo, M. C., *et al.* (2011). Introduction to the recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, 7(3), 257–262. Doi:10.1016/j.jalz.2011.03.004
- Jarrett, J. T., Berger, E. P., & Lansbury, P. T. (1993). 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(18), 4693–4697. Doi:10.1021/bi00069a001
- Jazi, G. A., Cauchi, J., Habib, A., Goyal, N., & Mozaffar, T. (2019). Primary Lateral Sclerosis (PLS) vs. Hereditary Spastic Paraparesis (HSP) in Presenilin-1 (PSEN1) related Dominant Dementia. Conference: American Academy of Neurology (AAN), 92 (15 Supplement) P4.4-030.
- Jobst, K. A., Smith, A. D., Szatmari, M., Jaskowski, A., King, E., Smith, A., *et al.* (1992). Detection in life of confirmed Alzheimer's disease using a simple measurement of medial temporal lobe atrophy by computed tomography. *The Lancet*, 340(8829), 1179–1183. Doi:10.1016/0140-6736(92)92890-r
- Jobst, K. A., Smith, A. D., Szatmari, M., Esiri, M. M., Jaskowski, A., Hindley, N., *et al.* (1994). Rapidly progressing atrophy of medial temporal lobe in Alzheimer's disease. *The Lancet*, 343(8901), 829–830. Doi:10.1016/s0140-6736(94)92028-1
- Joensen, P. (2011). Incidence of amyotrophic lateral sclerosis in the Faroe Islands. *Acta Neurologica Scandinavica*, 126(1), 62–66. Doi:10.1111/j.1600-0404.2011.01611.x
- Johnson-Kerner, B. L., Garcia Diaz, A., Ekins, S., & Wichterle, H. (2015). Kelch Domain of Gigaxonin Interacts with Intermediate Filament Proteins Affected in Giant Axonal Neuropathy. *PLOS ONE*, 10(10), e0140157. Doi:10.1371/journal.pone.0140157
- Justice, N. J. (2018). The relationship between stress and Alzheimer's disease. *Neurobiology of Stress*, 8, 127–133. Doi:10.1016/j.ynstr.2018.04.002
- Kametani, F., & Hasegawa, M. (2018). Reconsideration of Amyloid Hypothesis and Tau Hypothesis in Alzheimer's Disease. *Frontiers in Neuroscience*, 12. doi:10.3389/fnins.2018.00025

- Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alföldi, J., *et al.* (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*, 581(7809), 434–443. Doi:10.1038/s41586-020-2308-7
- Kato, J., Zhu, J., Liu, C., Stylianou, M., Hoffmann, V., Lizak, M. J., *et al.* (2011). ADP-Ribosylarginine Hydrolase Regulates Cell Proliferation and Tumorigenesis. *Cancer Research*, 71(15), 5327–5335. Doi:10.1158/0008-5472.can-10-0733
- Kato, J., Vekhter, D., Heath, J., Zhu, J., Barbieri, J. T., & Moss, J. (2015). Mutations of the functional *ARHI* allele in tumors from *ARHI* heterozygous mice and cells affect *ARHI* catalytic activity, cell proliferation and tumorigenesis. *Oncogenesis*, 4(6), e151–e151. Doi:10.1038/oncsis.2015.5
- Katzman, R. (1976). The Prevalence and Malignancy of Alzheimer Disease. *Archives of Neurology*, 33(4), 217. Doi:10.1001/archneur.1976.00500040001001
- Khera, R., & Das, N. (2009). Complement Receptor 1: Disease associations and therapeutic implications. *Molecular Immunology*, 46(5), 761–772. Doi:10.1016/j.molimm.2008.09.026
- Kishor, A., Fritz, S. E., & Hogg, J. R. (2019). Nonsense-mediated mRNA decay: the challenge of telling right from wrong in a complex transcriptome, *Wiley Interdiscip. Rev. RNA* 10(6), e1548, , <https://doi.org/10.1002/wrna.1548>
- Kisserli, A., Tabary, T., Cohen, J. H. M., Duret, V., & Mahmoudi, R. (2017). High-resolution Melting PCR for Complement Receptor 1 Length Polymorphism Genotyping: An Innovative Tool for Alzheimer’s Disease Gene Susceptibility Assessment. *Journal of Visualized Experiments*, (125). Doi:10.3791/56012
- Kita, M., Kuwata, Y., Murase, N., Akiyama, Y., & Usui, T. (2017). A Novel Truncation Mutation of the PRRT2 Gene Resulting in a 16-Amino-Acid Protein Causes Self-inducible Paroxysmal Kinesigenic Dyskinesia. *Movement Disorders Clinical Practice*, 4(4), 625–628. Doi:10.1002/mdc3.12500
- Knight, W. D., Kennedy, J., Mead, S., Rossor, M. N., Beck, J., Collinge, J., & Mummery, C. (2007). A novel presenilin 1 deletion (p.L166del) associated with early onset familial Alzheimer’s disease. *European Journal of Neurology*, 14(7), 829–831. Doi:10.1111/j.1468-1331.2007.01857.x
- Kobayashi, H., Kruger, R., Markopoulou, K., Wszolek, Z., Chase, B., Taka, H., *et al.*

- (2003). Haploinsufficiency at the α -synuclein gene underlies phenotypic severity in familial Parkinson's disease. *Brain*, 126(1), 32–42. doi:10.1093/brain/awg010
- Koch, P., Tamboli, I. Y., Mertens, J., Wunderlich, P., Ladewig, J., Stüber, K., *et al.* (2012). Presenilin-1 L166P Mutant Human Pluripotent Stem Cell–Derived Neurons Exhibit Partial Loss of γ -Secretase Activity in Endogenous Amyloid- β Generation. *The American Journal of Pathology*, 180(6), 2404–2416. doi:10.1016/j.ajpath.2012.02.012
- Kovalevich, J., & Langford, D. (2013). Considerations for the Use of SH-SY5Y Neuroblastoma Cells in Neurobiology. *Methods in Molecular Biology*, 9–21. Doi:10.1007/978-1-62703-640-5_2
- Kucukkilic, E., Brookes, K., Barber, I., Guetta-Baranes, T., Morgan, K., & Hollox, E. J. (2018). Complement receptor 1 gene (CR1) intragenic duplication and risk of Alzheimer's disease. *Human Genetics*, 137(4), 305–314. Doi:10.1007/s00439-018-1883-2
- Kuipers-Upmeijer, J. (2001). Primary lateral sclerosis: clinical, neurophysiological, and magnetic resonance findings. *Journal of Neurology, Neurosurgery & Psychiatry*, 71(5), 615–620. Doi:10.1136/jnnp.71.5.615
- Kumar, P., Henikoff, S., & Ng, P. C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nature Protocols*, 4(7), 1073–1081. doi:10.1038/nprot.2009.86
- Kure, S. (1892). An atypical case of Thomsen's disease. *Tokyo Igakukai Zasshi* 6, 505–514.
- Kurosaki, T., Popp, M. W., & Maquat, L. E. (2019). Quality and quantity control of gene expression by nonsense-mediated mRNA decay. *Nat. Rev. Mol. Cell. Biol.*, 20, 406–420. <https://doi.org/10.1038/s41580-019-0126-2>
- Kwart, D., Gregg, A., Scheckel, C., Murphy, E., Paquet, D., Duffield, M., *et al.* (2019). A Large Panel of Isogenic APP and PSEN1 Mutant Human iPSC Neurons Reveals Shared Endosomal Abnormalities Mediated by APP β -CTFs, Not A β . *Neuron*. doi:10.1016/j.neuron.2019.07.010
- Lacey, S., Chung, J., & Lin, H. (2014). A comparison of whole genome sequencing with exome sequencing for family-based association studies. *BMC Proceedings*, 8(1), S38. Doi:10.1186/1753-6561-8-s1-s38
- Lambert, J.-C., Heath, S., Even, G., Campion, D., Sleegers, K., *et al.* (2009). Genome-

- wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. *Nature Genetics*, 41(10), 1094–1099. Doi:10.1038/ng.439
- Landrum, M. J., Lee, J. M., Benson, M., Brown, G. R., Chao, C., Chitipiralla, S., *et al.* (2017). ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Research*, 46(D1), D1062–D1067. Doi:10.1093/nar/gkx1153
- Lee, H.-Y., Huang, Y., Bruneau, N., Roll, P., Roberson, E. D. O., Hermann, M., *et al.* (2012). Mutations in the Gene PRRT2 Cause Paroxysmal Kinesigenic Dyskinesia with Infantile Convulsions. *Cell Reports*, 1(1), 2–12. Doi:10.1016/j.celrep.2011.11.001
- Lévi-Strauss, M., & Mallat, M. (1987). Primary cultures of murine astrocytes produce C3 and factor B, two components of the alternative pathway of complement activation. *The Journal of Immunology*, 139(7), 2361–2366
- Li, M., Niu, F., Zhu, X., Wu, X., Shen, N., Peng, X., & Liu, Y. (2015). PRRT2 Mutant Leads to Dysfunction of Glutamate Signaling. *International Journal of Molecular Sciences*, 16(12), 9134–9151. Doi:10.3390/ijms16059134
- Li, N., Liu, K., Qiu, Y., Ren, Z., Dai, R., Deng, Y., & Qing, H. (2016). Effect of Presenilin Mutations on APP Cleavage; Insights into the Pathogenesis of FAD. *Frontiers in Aging Neuroscience*, 8. doi:10.3389/fnagi.2016.00051
- Li, Y., Song, D., Jiang, Y., Wang, J., Feng, R., Zhang, L., *et al.* (2015). CR1 rs3818361 Polymorphism Contributes to Alzheimer's Disease Susceptibility in Chinese Population. *Molecular Neurobiology*, 53(6), 4054–4059. Doi:10.1007/s12035-015-9343-7
- Lim, Y. Y., & Mormino, E. C. (2017). APOE genotype and early β -amyloid accumulation in older adults without dementia. *Neurology*, 89(10), 1028–1034. Doi:10.1212/wnl.0000000000004336
- Liu, D., Zhang, Y., Wang, Y., Chen, C., Li, X., Zhou, J., *et al.* (2016). Novel Locus for Paroxysmal Kinesigenic Dyskinesia Mapped to Chromosome 3q28-29. *Scientific Reports*, 6(1). Doi:10.1038/srep25790
- Liu, Q., Qi, Z., Wan, X.-H., Li, J.-Y., Shi, L., Lu, Q., *et al.* (2011). Mutations in PRRT2 result in paroxysmal dyskinesias with marked variability in clinical expression. *Journal of Medical Genetics*, 49(2), 79–82. Doi:10.1136/jmedgenet-2011-100653

- Liu, X. H., & Yu, X. C. (2019). Crystal structure of protein. PDB ID: 6IUX. *National Natural Science Foundation of China*. Doi:10.2210/pdb6IUX/pdb
- Liu, X.-R., Wu, M., He, N., Meng, H., Wen, L., Wang, J.-L., *et al.* (2012). Novel *PRRT2* mutations in paroxysmal dyskinesia patients with variant inheritance and phenotypes. *Genes, Brain and Behavior*, 12(2), 234–240. Doi:10.1111/gbb.12008
- Lobo, A., Lopez-Anton, R., Santabàrbara, J., de-la-Cámara, C., Ventura, T., Quintanilla, M. A., *et al.* (2011). Incidence and lifetime risk of dementia and Alzheimer's disease in a Southern European population. *Acta Psychiatrica Scandinavica*, 124(5), 372–383. Doi:10.1111/j.1600-0447.2011.01754.x
- Logroscino, G., Traynor, B. J., Hardiman, O., Chio, A., Mitchell, D., *et al.* J. (2009). Incidence of amyotrophic lateral sclerosis in Europe. *Journal of Neurology, Neurosurgery & Psychiatry*, 81(4), 385–390. Doi:10.1136/jnnp.2009.183525
- López-Pousa, S., Vilalta-Franch, J., Llinàs-Regla, J., Garre-Olmo, J., & Román, G. C. (2004). Incidence of Dementia in a Rural Community in Spain: The Girona Cohort Study. *Neuroepidemiology*, 23(4), 170–177. Doi:10.1159/000078502
- Lu, J. G., Bishop, J., Cheyette, S., Zhulin, I. B., Guo, S., Sobreira, N., & Brenner, S. E. (2017). A novel *PRRT2* pathogenic variant in a family with paroxysmal kinesigenic dyskinesia and benign familial infantile seizures. *Molecular Case Studies*, 4(1), a002287. Doi:10.1101/mcs.a002287
- Ludden, P., & Burris, R. (1976). Activating factor for the iron protein of nitrogenase from *Rhodospirillum rubrum*. *Science*, 194(4263), 424–426. Doi:10.1126/science.824729
- Ludolph, A., Drory, V., Hardiman, O., Nakano, I., Ravits, J., *et al.* (2015). A revision of the El Escorial criteria – 2015. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, 16(5-6), 291–292. Doi:10.3109/21678421.2015.1049183
- Luxenberg, J. S., May, C., Haxby, J. V., Grady, C., Moore, A., Berg, G., *et al.* (1987). Cerebral metabolism, anatomy, and cognition in monozygotic twins discordant for dementia of the Alzheimer type. *Journal of Neurology, Neurosurgery & Psychiatry*, 50(3), 333–340. Doi:10.1136/jnnp.50.3.333
- Lyoo, C. H., Cho, H., Choi, J. Y., Hwang, M. S., Hong, S. K., Kim, Y. J., *et al.* (2016). Tau Accumulation in Primary Motor Cortex of Variant Alzheimer's Disease

- with Spastic Paraparesis. *Journal of Alzheimer's Disease*, 51(3), 671–675.
Doi:10.3233/jad-151052
- Mahmoudi, R., Kisserli, A., Novella, J.-L., Donvito, B., Dramé, M., Réveil, B., *et al.* (2015). Alzheimer's disease is associated with low density of the long CR1 isoform. *Neurobiology of Aging*, 36(4), 1766.e5–1766.e12.
doi:10.1016/j.neurobiolaging.2015.01.006
- Maniatis, T., Fritsch, E. F., Sambrook J. (1982). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory
- Manubens, J. M., Martinez-Lage, J. M., Lacruz, F., Muruzabal, J., Larumbe, R., Guarch, C., *et al.* (1995). Prevalence of Alzheimer's disease and other dementing disorders in Pamplona: Spain. *Neuroepidemiology*, 14:155-64
- Marini, C., Conti, V., Mei, D., Battaglia, D., Lettori, D., Losito, E., *et al.* (2012). PRRT2 mutations in familial infantile seizures, paroxysmal dyskinesia, and hemiplegic migraine. *Neurology*, 79(21), 2109–2114. doi:10.1212/wnl.0b013e3182752ca2
- Marsh, J., & Alifragis, P. (2018). Synaptic dysfunction in Alzheimer's disease: the effects of amyloid beta on synaptic vesicle dynamics as a novel target for therapeutic intervention. *Neural regeneration research*, 13(4), 616–623.
<https://doi.org/10.4103/1673-5374.230276>
- Martin, L., Grigoryan, A., Wang, D., Wang, J., Breda, L., Rivella, S., *et al.* (2014). Identification and Characterization of Small Molecules That Inhibit Nonsense-Mediated RNA Decay and Suppress Nonsense p53 Mutations. *Cancer Research*, 74(11), 3104–3113. Doi:10.1158/0008-5472.can-13-2235
- Mashimo, M., Kato, J., & Moss, J. (2013). ADP-ribosyl-acceptor hydrolase 3 regulates poly (ADP-ribose) degradation and cell death during oxidative stress. *Proceedings of the National Academy of Sciences*, 110(47), 18964–18969.
Doi:10.1073/pnas.1312783110
- Mashimo, M., Kato, J., & Moss, J. (2014). Structure and function of the ARH family of ADP-ribosyl-acceptor hydrolases. *DNA Repair*, 23, 88–94.
Doi:10.1016/j.dnarep.2014.03.005
- Mashimo, M., Bu, X., Aoyama, K., Kato, J., Ishiwata-Endo, H., Stevens, L. A., Kasamatsu, A., Wolfe, L. A., Toro, C., Adams, D., *et al.* (2019). PARP1 inhibition alleviates injury in ARH3-deficient mice and human cells. *JCI Insight* 4, 124519. Doi:10.1172/jci.insight.124519

- Mayeux, R., & Stern, Y. (2012). Epidemiology of Alzheimer Disease. *Cold Spring Harbor Perspectives in Medicine*, 2(8), a006239–a006239. Doi:10.1101/cshperspect.a006239
- McGuire, S., Chanchani, S., & Khurana, D. S. (2018). Paroxysmal Dyskinesias. *Seminars in Pediatric Neurology*, 25, 75–81. Doi:10.1016/j.spnen.2017.12.007
- McGurk, L., Mojsilovic-Petrovic, J., Van Deerlin, V. M., Shorter, J., Kalb, R. G., Lee, V. M., *et al.* (2018). Nuclear poly(ADP-ribose) activity is a therapeutic target in amyotrophic lateral sclerosis. *Acta Neuropathologica Communications*, 6(1). Doi:10.1186/s40478-018-0586-1
- McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R., Kawas, C. H., *et al.* (2011). The diagnosis of dementia due to Alzheimer’s disease: Recommendations from the National Institute on Aging-Alzheimer’s Association workgroups on diagnostic guidelines for Alzheimer’s disease. *Alzheimer’s & Dementia*, 7(3), 263–269. Doi:10.1016/j.jalz.2011.03.005
- Medeiros, R., Baglietto-Vargas, D., & LaFerla, F. M. (2010). The Role of Tau in Alzheimer’s Disease and Related Disorders. *CNS Neuroscience & Therapeutics*, 17(5), 514–524. Doi:10.1111/j.1755-5949.2010.00177.x
- Mehrabian, S., Traykov, L., Rademakers, R., Jordanova, A., Cruts, M., Tournev, I., Raycheva, M., Dikova, C., Kremensky, I., & Van Broeckhoven, C. (2004) A novel *PSEN1* mutation in an EOAD family with spastic paraparesis and extrapyramidal signs. *Eur J Neurol*. 11(2), 16
- Méneret, A., Gaubebout, C., Riant, F., Vidailhet, M., Depienne, C., & Roze, E. (2013). PRRT2 mutations and paroxysmal disorders. *European Journal of Neurology*, 20(6), 872–878. Doi:10.1111/ene.12104
- Méneret, A., Grabli, D., Depienne, C., Gaubebout, C., Picard, F., Durr, A., *et al.* (2012). PRRT2 mutations: A major cause of paroxysmal kinesigenic dyskinesia in the European population. *Neurology*, 79(2), 170–174. Doi:10.1212/wnl.0b013e31825f06c3
- Mielke M. M. (2018). Sex and Gender Differences in Alzheimer’s Disease Dementia. *The Psychiatric times*, 35(11), 14–17.
- Moehlmann, T., Winkler, E., Xia, X., Edbauer, D., Murrell, J., Capell, A., *et al.* (2002). Presenilin-1 mutations of leucine 166 equally affect the generation of the Notch and APP intracellular domains independent of their effect on A 42 production.

- Proceedings of the National Academy of Sciences*, 99(12), 8025–8030.
Doi:10.1073/pnas.112686799
- Morgan, S., & Orrell, R. W. (2016). Pathogenesis of amyotrophic lateral sclerosis.
British Medical Bulletin, 119(1), 87–98. Doi:10.1093/bmb/ldw026
- Mort, M., Ivanov, D., Cooper, D. N., & Chuzhanova, N. A. (2008). A meta-analysis of nonsense mutations causing human genetic disease. *Human Mutation*, 29(8), 1037–1047. doi:10.1002/humu.20763
- Mortensen, S. A., Sander, B., Jensen, R. K., Pedersen, J. S., Golas, M. M., Jensenius, J. C., *et al.* (2017). Structure and activation of C1, the complex initiating the classical pathway of the complement cascade. *Proceedings of the National Academy of Sciences*, 114(5), 986–991. Doi:10.1073/pnas.1616998114
- Moss, J., Jacobson, M. K., & Stanley, S. J. (1985). Reversibility of arginine-specific mono(ADP-ribosylation): identification in erythrocytes of an ADP-ribose-L-arginine cleavage enzyme. *Proceedings of the National Academy of Sciences*, 82(17), 5603–5607. Doi:10.1073/pnas.82.17.5603
- Moss, J., Stanley, S. J., Nightingale, M. S., *et al.* (1992). Molecular and immunological characterization of ADP-ribosylarginine hydrolases. *J Biol Chem.*, 267(15), 10481-10488.
- Moss, M. L., White, J. M., Lambert, M. H., & Andrews, R. C. (2001). TACE and other ADAM proteases as targets for drug discovery. *Drug Discovery Today*, 6(8), 417–426. Doi:10.1016/s1359-6446(01)01738-x
- Moulton, P. V., & Yang, W. (2012). Air Pollution, Oxidative Stress, and Alzheimer's Disease. *Journal of Environmental and Public Health*, 2012, 1–9. Doi:10.1155/2012/472751
- Mu, W., Lu, H.-M., Chen, J., Li, S., & Elliott, A. M. (2016). Sanger Confirmation Is Required to Achieve Optimal Sensitivity and Specificity in Next-Generation Sequencing Panel Testing. *The Journal of Molecular Diagnostics*, 18(6), 923–932. Doi:10.1016/j.jmoldx.2016.07.006
- Mueller-Dieckmann, C., Kernstock, S., Lisurek, M., von Kries, J. P., Haag, F., Weiss, M. S., & Koch-Nolte, F. (2006). The structure of human ADP-ribosylhydrolase 3 (ARH3) provides insights into the reversibility of protein ADP-ribosylation. *Proceedings of the National Academy of Sciences*, 103(41), 15026–15031. Doi:10.1073/pnas.0606762103

- Nelson, O., Tu, H., Lei, T., Bentahir, M., de Strooper, B., & Bezprozvanny, I. (2007). Familial Alzheimer disease-linked mutations specifically disrupt Ca²⁺ leak function of presenilin 1. *J Clin Invest.* 117(5), 1230-1239. <https://doi.org/10.1172/JCI30447>
- Nguyen, H. P., Van Broeckhoven, C., & van der Zee, J. (2018). ALS Genes in the Genomic Era and their Implications for FTD. *Trends in Genetics*, 34(6), 404–423. Doi:10.1016/j.tig.2018.03.001
- Niere, M., Mashimo, M., Agledal, L., Dölle, C., Kasamatsu, A., Kato, J., *et al.* (2012). ADP-ribosylhydrolase 3 (ARH3), Not Poly(ADP-ribose) Glycohydrolase (PARG) Isoforms, Is Responsible for Degradation of Mitochondrial Matrix-associated Poly(ADP-ribose). *Journal of Biological Chemistry*, 287(20), 16088–16102. Doi:10.1074/jbc.m112.349183
- Niu, H., Álvarez-Álvarez, I., Guillén-Grima, F., & Aguinaga-Ontoso, I. (2017). Prevalencia e incidencia de la enfermedad de Alzheimer en Europa: metaanálisis. *Neurología*, 32(8), 523–532. Doi:10.1016/j.nrl.2016.02.016
- Noris, M., & Remuzzi, G. (2013). Overview of Complement Activation and Regulation. *Seminars in Nephrology*, 33(6), 479–492. Doi:10.1016/j.semnephrol.2013.08.001
- Nussbaum, R. L., & Ellis, C. E. (2003). Alzheimer's Disease and Parkinson's Disease. *New England Journal of Medicine*, 348(14), 1356–1364. Doi:10.1056/nejm2003ra020003
- O'Leary, N. A., Wright, M. W., Brister, J. R., Ciufu, S., Haddad, D., McVeigh, R., *et al.* (2015). Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Research*, 44(D1), D733–D745. Doi:10.1093/nar/gkv1189
- Ohm, T. G., Kirca, M., Bohl, J., Scharnagl, H., Groß, W., & März, W. (1995). Apolipoprotein E polymorphism influences not only cerebral senile plaque load but also Alzheimer-type neurofibrillary tangle formation. *Neuroscience*, 66(3), 583–587. Doi:10.1016/0306-4522(94)00596-w
- Oka, S., Kato, J., & Moss, J. (2005). Identification and Characterization of a Mammalian 39-kDa Poly(ADP-ribose) Glycohydrolase. *Journal of Biological Chemistry*, 281(2), 705–713. Doi:10.1074/jbc.m510290200
- Ono, S., Yoshiura, K., Kinoshita, A., Kikuchi, T., Nakane, Y., Kato, N., *et al.* (2012).

- Mutations in PRRT2 responsible for paroxysmal kinesigenic dyskinesias also cause benign familial infantile convulsions. *Journal of Human Genetics*, 57(5), 338–341. Doi:10.1038/jhg.2012.23
- Ono, T., Kasamatsu, A., Oka, S., & Moss, J. (2006). The 39-kDa poly(ADP-ribose) glycohydrolase ARH3 hydrolyzes O-acetyl-ADP-ribose, a product of the Sir2 family of acetyl-histone deacetylases. *Proceedings of the National Academy of Sciences*, 103(45), 16687–16691. Doi:10.1073/pnas.0607911103
- Oskarsson, B., Horton, D. K., & Mitsumoto, H. (2015). Potential Environmental Factors in Amyotrophic Lateral Sclerosis. *Neurologic Clinics*, 33(4), 877–888. Doi:10.1016/j.ncl.2015.07.009
- Palau-Bargues, J., Rubio-Agustí, I., Burguera, J. A., Vílchez-Padilla, J. J., & Villanueva, V. (2010). Familial idiopathic paroxysmal kinesigenic dyskinesia: its natural history and a descriptive study in three Spanish families. *Neurol*, 50, 390-394 <https://doi.org/10.33588/rn.5007.2009421>
- Palazzo, L., Mikolčević, P., Mikoč, A., & Ahel, I. (2019). ADP-ribosylation signalling and human disease. *Open Biology*, 9(4), 190041. doi:10.1098/rsob.190041
- Palop, J. J., & Mucke, L. (2010). Amyloid- β -induced neuronal dysfunction in Alzheimer's disease: from synapses toward neural networks. *Nature Neuroscience*, 13(7), 812–818. Doi:10.1038/nn.2583
- Panas, M., Karadima, G., Kalfakis, N., Psarrou, O., Floroskoufi, P., Kladi, A., *et al.* (2000). Genotyping of presenilin-1 polymorphism in amyotrophic lateral sclerosis. *Journal of Neurology*, 247(12), 940–942. doi:10.1007/s004150070050
- Pantieri, R., Pardini, M., Cecconi, M., Dagna-Bricarelli, F., Vitali, A., Piccini, A., *et al.* (2005). A novel presenilin 1 L166H mutation in a pseudo-sporadic case of early-onset Alzheimer's disease. *Neurological Sciences*, 26(5), 349–350. Doi:10.1007/s10072-005-0499-1
- Parihar, M. S., & Brewer, G. J. (2010). Amyloid- β as a Modulator of Synaptic Plasticity. *Journal of Alzheimer's Disease*, 22(3), 741–763. Doi:10.3233/jad-2010-101020
- Pasinelli, P., & Brown, R. H. (2006). Molecular biology of amyotrophic lateral sclerosis: insights from genetics. *Nature Reviews Neuroscience*, 7(9), 710–723. Doi:10.1038/nrn1971
- Pegueroles J., Jiménez A., Vilaplana E., Montal V., Carmona-Iragui M., Pané A., Alcolea D., Videla L., Casajoana A., Clarimón J., Ortega E., Vidal J., Blesa R.,

- et al.* (2018). Obesity and Alzheimer's disease, does the obesity paradox really exist? A magnetic resonance imaging study. *Oncotarget*, 9, 34691–34698. Doi: 10.18632/oncotarget.26162
- Pekny, M., & Pekna, M. (2014). Astrocyte Reactivity and Reactive Astrogliosis: Costs and Benefits. *Physiological Reviews*, 94(4), 1077–1098. Doi:10.1152/physrev.00041.2013
- Peters, R., Ee, N., Peters, J., Booth, A., Mudway, I., & Anstey, K. J. (2019). Air Pollution and Dementia: A Systematic Review. *Journal of Alzheimer's Disease*, 1–19. Doi:10.3233/jad-180631
- Pilotto, A., Padovani, A., & Borroni, B. (2013). Clinical, Biological, and Imaging Features of Monogenic Alzheimer's Disease. *BioMed Research International*, 2013, 1–9. Doi:10.1155/2013/689591
- Prabhakara, S., & Anbazhagan, K. (2014). Molecular analysis of PRRT2 gene in a case of paroxysmal kinesigenic dyskinesia patient. *Annals of Indian Academy of Neurology*, 17(4), 459. Doi:10.4103/0972-2327.144039
- Pradas, J., Puig, T., Rojas-García, R., Viguera, M. L., Gich, I., & Logroscino, G. (2013). Amyotrophic lateral sclerosis in Catalonia: A population based study. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, 14(4), 278–283. Doi:10.3109/21678421.2012.749915
- Praline, J., Guennoc, A.-M., Vourc'h, P., De Toffol, B., & Corcia, P. (2010). Primary lateral sclerosis may occur within familial amyotrophic lateral sclerosis pedigrees. *Amyotrophic Lateral Sclerosis*, 11(1-2), 154–156. Doi:10.3109/17482960802483038
- Pringle, C. E., Hudson, A. J., Munoz, D. G., Kiernan, J. A., Brown, W. F., & Ebers, G. C. (1992). Primary lateral sclerosis: clinical features, neuropathology and diagnostic criteria. *Brain*, 115(2), 495–520. Doi:10.1093/brain/115.2.495
- R Core Team (2019). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria. URL <https://www.R-project.org/>
- Raber, J., Huang, Y., & Ashford, J. W. (2004). ApoE genotype accounts for the vast majority of AD risk and AD pathology. *Neurobiology of Aging*, 25(5), 641–650. Doi:10.1016/j.neurobiolaging.2003.12.023
- Rack, J. G. M., Ariza, A., Drown, B. S., Henfrey, C., Bartlett, E., Shirai, T., *et al.*

- (2018). (ADP-ribosyl)hydrolases: Structural Basis for Differential Substrate Recognition and Inhibition. *Cell Chemical Biology*.
Doi:10.1016/j.chembiol.2018.11.001
- Rack, J. G. M., Palazzo, L., & Ahel, I. (2020). (ADP-ribosyl)hydrolases: structure, function, and biology. *Genes & Development*. Doi:10.1101/gad.334631.119
- Raghavan, N. S., Brickman, A. M., Andrews, H., Manly, J. J., Schupf, N., *et al.* (2018). Whole-exome sequencing in 20,197 persons for rare variants in Alzheimer's disease. *Annals of Clinical and Translational Neurology*, 5(7), 832–842. Doi:10.1002/acn3.582
- Ramakers, C., Ruijter, J. M., Deprez, R. H. L., & Moorman, A. F. (2003). Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neuroscience Letters*, 339(1), 62–66. Doi:10.1016/s0304-3940(02)01423-4
- Ramos-Cejudo, J., Wisniewski, T., Marmar, C., Zetterberg, H., Blennow, K., de Leon, M. J., & Fossati, S. (2018). Traumatic Brain Injury and Alzheimer's Disease: The Cerebrovascular Link. *EbioMedicine*, 28, 21–30. Doi:10.1016/j.ebiom.2018.01.021
- Rao, V. S., Cupples, A., van Duijn, C. M., Kurz, A., Green, R. C., Chui, H., Duara, R., Auerbach, S. A., Volicer, L., Wells, J., van Broeckhoven, C., Growdon, J. H., Haines, J. L., & Farrer, L. A. (1996). Evidence for major gene inheritance of Alzheimer disease in families of patients with and without apolipoprotein E epsilon 4. *Am J Hum Genet*, 59(3):664-75
- Rice, D. P., Fillit, H. M., Max, W., Knopman, D. S., Lloyd, J. R., & Duttgupta, S. (2001). Prevalence, costs, and treatment of Alzheimer's disease and related dementia: A managed care perspective. *Am. J. Manag. Care* 7, 809–818.
- Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., *et al.* (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in Medicine*, 17(5), 405–423. Doi:10.1038/gim.2015.30
- Richens, J. L., Spencer, H. L., Butler, M., Cantlay, F., Vere, K.-A., Bajaj, N., *et al.* (2016). Rationalising the role of Keratin 9 as a biomarker for Alzheimer's disease. *Scientific Reports*, 6(1). Doi:10.1038/srep22962

- Roses, M.D, A. D. (1996). Apolipoprotein E alleles as risk factors in Alzheimer's disease. *Annual Review of Medicine*, 47(1), 387–400. Doi:10.1146/annurev.med.47.1.387
- Rossi, P., Sterlini, B., Castroflorio, E., Marte, A., Onofri, F., Valtorta, F., *et al.* (2018). A novel topology of proline-rich transmembrane protein 2 (PRRT2): Hints for an intracellular function at the synapse. *Journal of Biological Chemistry*, 293(12), 4581–4581. Doi:10.1074/jbc.aac118.002651
- Roze, E., Méneret, A., & Vidailhet, M. (2015). Paroxysmal Movement Disorders. *Movement Disorders*, 767–778. Doi:10.1016/b978-0-12-405195-9.00048-2
- Ruijter, J. M., Pfaffl, M. W., Zhao, S., Spiess, A. N., Boggy, G., Blom, J., *et al.* (2013). Evaluation of qPCR curve analysis methods for reliable biomarker discovery: Bias, resolution, precision, and implications. *Methods*, 59(1), 32–46. Doi:10.1016/j.ymeth.2012.08.011
- Salzberg, S. L. (2018). Open questions: How many genes do we have? *BMC Biology*, 16(1). doi:10.1186/s12915-018-0564-x
- Sando, S. B., Melquist, S., Cannon, A., Hutton, M. L., Sletvold, O., Saltvedt, I., *et al.* (2008). APOE ε4 lowers age at onset and is a high risk factor for Alzheimer's disease; A case control study from central Norway. *BMC Neurology*, 8(1). Doi:10.1186/1471-2377-8-9
- Santos-Mandujano, R. A., Ryan, N. S., Chávez-Gutiérrez, L., Sánchez-Torres, C., & Meraz-Ríos, M. A. (2019). Clinical Association of White Matter Hyperintensities Localization in a Mexican Family with Spastic Paraparesis Carrying the PSEN1 A431E Mutation. *Journal of Alzheimer's Disease*, 1–9. doi:10.3233/jad-190978
- Sassi, C., Guerreiro, R., Gibbs, R., Ding, J., Lupton, M. K., Troakes, C., *et al.* (2014). Exome sequencing identifies 2 novel presenilin 1 mutations (p.L166V and p.S230R) in British early-onset Alzheimer's disease. *Neurobiology of Aging*, 35(10), 2422.e13–2422.e16. doi:10.1016/j.neurobiolaging.2014.04.026
- Scahill, R. I., Schott, J. M., Stevens, J. M., Rossor, M. N., & Fox, N. C. (2002). Mapping the evolution of regional atrophy in Alzheimer's disease: Unbiased analysis of fluid-registered serial MRI. *Proceedings of the National Academy of Sciences*, 99(7), 4703–4707. Doi:10.1073/pnas.052587399
- Schelosky, L. D. (2010). Paroxysmal dyskinesias. *The Paroxysmal Disorders*, 113–

129. Doi:10.1017/cbo9780511781780.009
- Scheltens, P., Blennow, K., Breteler, M. M. B., de Strooper, B., Frisoni, G. B., Salloway, S., & Van der Flier, W. M. (2016). Alzheimer's disease. *The Lancet*, 388(10043), 505–517. Doi:10.1016/s0140-6736(15)01124-1
- Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, 9(7), 671–675.
Doi:10.1038/nmeth.2089
- Schröder, B., Franz, B., Hempfling, P., Selbert, M., Jürgens, T., Kretzschmar, H. A., *et al.* (2001). Polymorphisms within the prion-like protein gene (Prnd) and their implications in human prion diseases, Alzheimer's disease and other neurological disorders. *Human Genetics*, 109(3), 319–325.
Doi:10.1007/s004390100591
- Schubert, J., Paravidino, R., Becker, F., Berger, A., Bebek, N., Bianchi, A., *et al.* (2012). PRRT2 Mutations are the major cause of benign familial infantile seizures. *Human Mutation*, 33(10), 1439–1443. Doi:10.1002/humu.22126
- Schweingruber, C., Rufener, S. C., Zünd, D., Yamashita, A., & Mühlemann, O. (2013). Nonsense-mediated mRNA decay — Mechanisms of substrate mRNA recognition and degradation in mammalian cells. *Biochimica et Biophysica Acta (BBA) – Gene Regulatory Mechanisms*, 1829(6-7), 612–623.
Doi:10.1016/j.bbagr.2013.02.005
- Schwenk, J., Baehrens, D., Haupt, A., Bildl, W., Boudkkazi, S., Roeper, J., *et al.* (2014). Regional Diversity and Developmental Dynamics of the AMPA-Receptor Proteome in the Mammalian Brain. *Neuron*, 84(1), 41–54.
Doi:10.1016/j.neuron.2014.08.044
- Sehna, D., Deshpande, M., Vařeková, R. S., Mir, S., Berka, K., Midlik, A., *et al.* (2017). LiteMol suite: interactive web-based visualization of large-scale macromolecular structure data. *Nature Methods*, 14(12), 1121–1122.
Doi:10.1038/nmeth.4499
- Selkoe, D. J. (1991). The molecular pathology of Alzheimer's disease. *Neuron*, 6(4), 487–498. Doi:10.1016/0896-6273(91)90052-2
- Sennvik, K., Fastbom, J., Blomberg, M., Wahlund, L.-O., Winblad, B., & Benedikz, E. (2000). Levels of α - and β -secretase cleaved amyloid precursor protein in the

- cerebrospinal fluid of Alzheimer's disease patients. *Neuroscience Letters*, 278(3), 169–172. Doi:10.1016/s0304-3940(99)00929-5
- Sharifi, R., Morra, R., Denise Appel, C., Tallis, M., Chioza, B., Jankevicius, G., *et al.* (2013). Deficiency of terminal ADP-ribose protein glycohydrolase TARG1/C6orf130 in neurodegenerative disease. *The EMBO Journal*, 32(9), 1225–1237. Doi:10.1038/emboj.2013.51
- Sharp, E. S., & Gatz, M. (2011). Relationship Between Education and Dementia. *Alzheimer Disease & Associated Disorders*, 25(4), 289–304. Doi:10.1097/wad.0b013e318211c83c
- Shaw, P. J., & Ince, P. G. (1997). Glutamate, excitotoxicity and amyotrophic lateral sclerosis. *Journal of Neurology*, 244(S2), S3–S14. Doi:10.1007/bf03160574
- Shen, X., Ellis, R. E., Lee, K., Liu, C.-Y., Yang, K., Solomon, A., *et al.* (2001). Complementary Signaling Pathways Regulate the Unfolded Protein Response and Are Required for *C. elegans* Development. *Cell*, 107(7), 893–903. Doi:10.1016/s0092-8674(01)00612-2
- Shen, Y., Yang, L., & Li, R. (2013). What does complement do in Alzheimer's disease? Old molecules with new insights. *Translational Neurodegeneration*, 2(1), 21. Doi:10.1186/2047-9158-2-21
- Shi, Y., Lin, S., Staats, K. A., Li, Y., Chang, W.-H., Hung, S.-T., *et al.* (2018). Haploinsufficiency leads to neurodegeneration in C9ORF72 ALS/FTD human induced motor neurons. *Nature Medicine*, 24(3), 313–325. doi:10.1038/nm.4490
- Singer, H. S., Mink, J. W., Gilbert, D. L., & Jankovic, J. (2010). Paroxysmal Dyskinesias. *Movement Disorders in Childhood*, 66–74. Doi:10.1016/b978-0-7506-9852-8.00008-4
- Singhrai, S. K., Neal, J. W., Rushmere, N. K., Morgan, B. P., & Gasque, P. (2000). Spontaneous Classical Pathway Activation and Deficiency of Membrane Regulators Render Human Neurons Susceptible to Complement Lysis. *The American Journal of Pathology*, 157(3), 905–918. Doi:10.1016/s0002-9440(10)64604-4
- Small, G. W. (2000). Early diagnosis of Alzheimer's disease: update on combining genetic and brain-imaging measures. *Dialogues Clin Neurosci.* 2(3), 241–246

- Smith, A. D. (2002). Imaging the progression of Alzheimer pathology through the brain. *Proceedings of the National Academy of Sciences*, 99(7), 4135–4137. Doi:10.1073/pnas.082107399
- Smith, E. F., Shaw, P. J., & De Vos, K. J. (2017). The role of mitochondria in amyotrophic lateral sclerosis. *Neuroscience Letters*. Doi:10.1016/j.neulet.2017.06.052
- Smith, S. J., Towers, N., Saldanha, J. W., Shang, C. A., Mahmood, S. R., Taylor, W. R., & Mohun, T. J. (2016). The cardiac-restricted protein ADP-ribosylhydrolase-like 1 is essential for heart chamber outgrowth and acts on muscle actin filament assembly. *Developmental Biology*, 416(2), 373–388. Doi:10.1016/j.ydbio.2016.05.006
- Smith, S. J., Towers, N., Demetriou, K., & Mohun, T. J. (2020). Defective heart chamber growth and myofibrillogenesis after knockout of adprh1 gene function by targeted disruption of the ancestral catalytic active site. *PLOS ONE*, 15(7), e0235433. Doi:10.1371/journal.pone.0235433
- Solomon, A., Kivipelto, M., Wolozin, B., Zhou, J., & Whitmer, R. A. (2009). Midlife Serum Cholesterol and Increased Risk of Alzheimer's and Vascular Dementia Three Decades Later. *Dementia and Geriatric Cognitive Disorders*, 28(1), 75–80. Doi:10.1159/000231980
- Sperling, R. A., Aisen, P. S., Beckett, L. A., Bennett, D. A., Craft, S., Fagan, A. M., *et al.* (2011). Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, 7(3), 280–292. Doi:10.1016/j.jalz.2011.03.003
- Spiller, W. C. (1905a). Primary degeneration of the pyramidal tracts: a study of 8 cases with necropsy. *Univ Pa Med Bull* 1905;17:390–5
- Spiller, W. C. (1905b). Primary degeneration of the pyramidal tracts: a study of 8 cases with necropsy. *Univ Pa Med Bull* 1905;17:407–14
- Sreedharan, J., Blair, I. P., Tripathi, V. B., Hu, X., Vance, C., Rogelj, B., *et al.* (2008). TDP-43 Mutations in Familial and Sporadic Amyotrophic Lateral Sclerosis. *Science*, 319(5870), 1668–1672. Doi:10.1126/science.1154584
- Stelzer, G., Rosen, N., Plaschkes, I., Zimmerman, S., Twik, M., Fishilevich, S., *et al.* (2016). The GeneCards Suite: From Gene Data Mining to Disease Genome

- Sequence Analyses. *Current Protocols in Bioinformatics*, 1.30.1–1.30.33.
doi:10.1002/cpbi.5
- Stelzl, U., Worm, U., Lalowski, M., Haenig, C., Brembeck, F. H., Goehler, H., *et al.* (2005). A Human Protein-Protein Interaction Network: A Resource for Annotating the Proteome. *Cell*, 122(6), 957–968.
Doi:10.1016/j.cell.2005.08.029
- Strittmatter, W. J., Saunders, A. M., Schmechel, D., Pericak-Vance, M., Enghild, J., Salvesen, G. S., & Roses, A. D. (1993). Apolipoprotein E: high-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease. *Proceedings of the National Academy of Sciences*, 90(5), 1977–1981. Doi:10.1073/pnas.90.5.1977
- Strong, M. J., Lomen-Hoerth, C., Caselli, R. J., Bigio, E. H., & Yang, W. (2003). Cognitive impairment, frontotemporal dementia, and the motor neuron diseases. *Ann. Neurol.*, 54: S20-23
- Suryadi, K., & Shine, N. (2011). HPLC Method to Assay the ADP – Ribosyltransferase Activity of Cholera Toxin. List Labs POSTER – Presented at the 14th Annual Bay Area Microbial Pathogenesis Symposium (BAMPS) Meeting in San Francisco, California.
- Szaruga, M., Munteanu, B., Lismont, S., Veugelen, S., Horr , K., Mercken, M., *et al.* (2017). Alzheimer’s-Causing Mutations Shift A β Length by Destabilizing γ -Secretase-A β n Interactions. *Cell*, 170(3), 443–456.e14.
doi:10.1016/j.cell.2017.07.004
- Tan, G.-H., Liu, Y.-Y., Wang, L., Li, K., Zhang, Z.-Q., Li, H.-F., *et al.* (2017). PRRT2 deficiency induces paroxysmal kinesigenic dyskinesia by regulating synaptic transmission in cerebellum. *Cell Research*, 28(1), 90–110.
Doi:10.1038/cr.2017.128
- Tanzi, R. E., & Bertram, L. (2005). Twenty Years of the Alzheimer’s Disease Amyloid Hypothesis: A Genetic Perspective. *Cell*, 120(4), 545–555.
Doi:10.1016/j.cell.2005.02.008
- Tanzi, R. E. (2012). The Genetics of Alzheimer Disease. *Cold Spring Harbor Perspectives in Medicine*, 2(10), a006296–a006296.
Doi:10.1101/cshperspect.a006296
- Tartaglia, M. C., Rowe, A., Findlater, K., Orange, J. B., Grace, G., & Strong, M. J.

- (2007). Differentiation Between Primary Lateral Sclerosis and Amyotrophic Lateral Sclerosis. *Archives of Neurology*, 64(2), 232.
Doi:10.1001/archneur.64.2.232
- Terry, R. D., & Davies, P. (1980). Dementia of the Alzheimer Type. *Annual Review of Neuroscience*, 3(1), 77–95. Doi:10.1146/annurev.ne.03.030180.000453
- Tola-Arribas, M. A., Yugueros, M. I., Garea, M. J., Ortega-Valín, F., Cerón-Fernández, A., Fernández-Malvido, B., *et al.* (2013). Prevalence of Dementia and Subtypes in Valladolid, Northwestern Spain: The DEMINVALL Study. *PLoS ONE*, 8(10), e77688. Doi:10.1371/journal.pone.0077688
- Tsai, M., Nian, F., Hsu, M., Liu, W., Liu, Y., Liu, C., *et al.* (2019). *PRRT2* missense mutations cluster near C-terminus and frequently lead to protein mislocalization. *Epilepsia*. Doi:10.1111/epi.14725
- Tu, H., Nelson, O., Bezprozvanny, A., Wang, Z., Lee, S.-F., Hao, Y.-H., *et al.* (2006). Presenilins Form ER Ca²⁺ Leak Channels, a Function Disrupted by Familial Alzheimer's Disease-Linked Mutations. *Cell*, 126(5), 981–993. doi:10.1016/j.cell.2006.06.059
- Tuomi, J. M., Voorbraak, F., Jones, D. L., & Ruijter, J. M. (2010). Bias in the Cq value observed with hydrolysis probe based quantitative PCR can be corrected with the estimated PCR efficiency value. *Methods*, 50(4), 313–322.
Doi:10.1016/j.ymeth.2010.02.003
- Ugozzoli, L., & Wallace, R. (1991). Allele-specific polymerase chain reaction. *Methods*, 2(1), 42–48. Doi:10.1016/s1046-2023(05)80124-0
- Uhlen, M., Fagerberg, L., Hallstrom, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., *et al.* (2015). Tissue-based map of the human proteome. *Science*, 347(6220), 1260419–1260419. Doi:10.1126/science.1260419
- Ulep, M. G., Saraon, S. K., & McLea, S. (2018). Alzheimer Disease. *The Journal for Nurse Practitioners*, 14(3), 129–135. Doi:10.1016/j.nurpra.2017.10.014
- Valdmanis, P. N., Dupré, N., & Rouleau, G. A. (2008). A Locus for Primary Lateral Sclerosis on Chromosome 4ptel-4p16.1. *Archives of Neurology*, 65(3).
Doi:10.1001/archneur.65.3.383
- Valente, P., Castroflorio, E., Rossi, P., Fadda, M., Sterlini, B., Cervigni, R. I., *et al.* (2016). *PRRT2* Is a Key Component of the Ca²⁺ -Dependent Neurotransmitter Release Machinery. *Cell Reports*, 15(1), 117–131.

Doi:10.1016/j.celrep.2016.03.005

- Valente, P., Romei, A., Fadda, M., Sterlini, B., Lonardoni, D., Forte, N., *et al.* (2018). Constitutive Inactivation of the PRRT2 Gene Alters Short-Term Synaptic Plasticity and Promotes Network Hyperexcitability in Hippocampal Neurons. *Cerebral Cortex*. Doi:10.1093/cercor/bhy079
- Valtorta, F., Benfenati, F., Zara, F., & Meldolesi, J. (2016). PRRT2: from Paroxysmal Disorders to Regulation of Synaptic Function. *Trends in Neurosciences*, 39(10), 668–679. Doi:10.1016/j.tins.2016.08.005
- Van Blitterswijk, M., & Landers, J. E. (2010). RNA processing pathways in amyotrophic lateral sclerosis. *Neurogenetics*, 11(3), 275–290. Doi:10.1007/s10048-010-0239-4
- Van Cauwenberghe, C., Van Broeckhoven, C., & Sleegers, K. (2015). The genetic landscape of Alzheimer disease: clinical implications and perspectives. *Genetics in Medicine*, 18(5), 421–430. Doi:10.1038/gim.2015.117
- Van Dijk, E. L., Auger, H., Jaszczyszyn, Y., & Thermes, C. (2014). Ten years of next-generation sequencing technology. *Trends in Genetics*, 30(9), 418–426. Doi:10.1016/j.tig.2014.07.001
- Van Eerdewegh, P., Little, R. D., Dupuis, J., Del Mastro, R. G., Falls, K., Simon, J., *et al.* (2002). Association of the ADAM33 gene with asthma and bronchial hyperresponsiveness. *Nature*, 418(6896), 426–430. Doi:10.1038/nature00878
- Van Mossevelde, S., van der Zee, J., Gijssels, I., Engelborghs, S., Sieben, A., Van Langenhove, T., *et al.* (2015). Clinical features of *TBKI* carriers compared with *C9orf72*, *GRN* and non-mutation carriers in a Belgian cohort. *Brain*, 139(2), 452–467. Doi:10.1093/brain/awv358
- Van Rheenen, W., van Blitterswijk, M., Huisman, M. H. B., Vlam, L., van Doormaal, P. T. C., Seelen, M., *et al.* (2012). Hexanucleotide repeat expansions in *C9ORF72* in the spectrum of motor neuron diseases. *Neurology*, 79(9), 878–882. Doi:10.1212/wnl.0b013e3182661d14
- Van Vliet, R., Breedveld, G., de Rijk-van Andel, J., Brilstra, E., Verbeek, N., Verschuuren-Bemelmans, C., *et al.* (2012). PRRT2 phenotypes and penetrance of paroxysmal kinesigenic dyskinesia and infantile convulsions. *Neurology*, 79(8), 777–784. doi:10.1212/wnl.0b013e3182661fe3
- Vance, C., Rogelj, B., Hortobagyi, T., De Vos, K. J., Nishimura, A. L., Sreedharan, J.,

- et al.* (2009). Mutations in FUS, an RNA Processing Protein, Cause Familial Amyotrophic Lateral Sclerosis Type 6. *Science*, 323(5918), 1208–1211.
Doi:10.1126/science.1165942
- Venselaar, H., te Beek, T. A., Kuipers, R. K., Hekkelman, M. L., & Vriend, G. (2010). Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. *BMC Bioinformatics*, 11(1), 548. Doi:10.1186/1471-2105-11-548
- Wainaina, M. N., Chen, Z., & Zhong, C. (2014). Environmental factors in the development and progression of late-onset Alzheimer's disease. *Neuroscience Bulletin*, 30(2), 253–270. Doi:10.1007/s12264-013-1425-9
- Wang, H.-X., Li, H.-F., Liu, G.-L., Wen, X.-D., & Wu, Z.-Y. (2016). Mutation Analysis of MR-1, SLC2A1, and CLCN1 in 28 PRRT2-negative Paroxysmal Kinesigenic Dyskinesia Patients. *Chinese Medical Journal*, 129(9), 1017–1021. Doi:10.4103/0366-6999.180529
- Wang, J.-L., Cao, L., Li, X.-H., Hu, Z.-M., Li, J.-D., Zhang, J.-G., *et al.* (2011). Identification of PRRT2 as the causative gene of paroxysmal kinesigenic dyskinesias. *Brain*, 134(12), 3493–3501. Doi:10.1093/brain/awr289
- Wang, Z., & Moulton, J. (2001). SNPs, protein structure, and disease. *Human Mutation*, 17(4), 263–270. doi:10.1002/humu.22
- Ward, M. M. (2013). Estimating Disease Prevalence and Incidence Using Administrative Data: Some Assembly Required. *The Journal of Rheumatology*, 40(8), 1241–1243. Doi:10.3899/jrheum.130675
- Whittington, R., Papon, M.-A., Chouinard-Decorte, F., & Planel, E. (2010). Hypothermia and Alzheimers Disease Neuropathogenic Pathways. *Current Alzheimer Research*, 7(8), 717–725. Doi:10.2174/156720510793611646
- Williamson, M. P. (1994). The structure and function of proline-rich regions in proteins. *Biochemical Journal*, 297(2), 249–260. doi:10.1042/bj2970249
- Witteveen, J. S., Willemsen, M. H., Dombroski, T. C. D., van Bakel, N. H. M., Nillesen, W. M., van Hulst, J. A., *et al.* (2016). Haploinsufficiency of MeCP2-interacting transcriptional co-repressor SIN3A causes mild intellectual disability by affecting the development of cortical integrity. *Nature Genetics*, 48(8), 877–887. Doi:10.1038/ng.3619
- Wu, L., Tang, H.-D., Huang, X.-J., Zheng, L., Liu, X.-L., Wang, T., *et al.* (2014).

- PRRT2 truncated mutations lead to nonsense-mediated mRNA decay in Paroxysmal Kinesigenic Dyskinesia. *Parkinsonism & Related Disorders*, 20(12), 1399–1404. Doi:10.1016/j.parkreldis.2014.10.012
- Xu, Z., Cork, L. C., Griffin, J. W., & Cleveland, D. W. (1993). Increased expression of neurofilament subunit NF-L produces morphological alterations that resemble the pathology of human motor neuron disease. *Cell*, 73(1), 23–33. Doi:10.1016/0092-8674(93)90157-1
- Yamaguchi-Kabata, Y., Shimada, M. K., Hayakawa, Y., Minoshima, S., Chakraborty, R., Gojobori, T., & Imanishi, T. (2008). Distribution and Effects of Nonsense Polymorphisms in Human Genes. *PLoS ONE*, 3(10), e3393. doi:10.1371/journal.pone.0003393
- Yan, D., Zhang, Y., Liu, L., & Yan, H. (2016). Pesticide exposure and risk of Alzheimer's disease: a systematic review and meta-analysis. *Scientific Reports*, 6(1). Doi:10.1038/srep32222
- Yanagisawa, M., Planel, E., Ishiguro, K., & Fujita, S. C. (1999). Starvation induces tau hyperphosphorylation in mouse brain: implications for Alzheimer's disease. *FEBS Letters*, 461(3), 329–333. Doi:10.1016/s0014-5793(99)01480-5
- Yang, X., Zhang, Y., Xu, X., Wang, S., Yang, Z., Wu, Y., *et al.* (2013). Phenotypes and PRRT2 mutations in Chinese families with benign familial infantile epilepsy and infantile convulsions with paroxysmal choreoathetosis. *BMC Neurology*, 13(1). doi:10.1186/1471-2377-13-209
- Yang, Y., Zhang, L., Lynch, D. R., Lukas, T., Ahmeti, K., Sleiman, P. M. A., *et al.* (2016). Compound heterozygote mutations in *SPG7* in a family with adult-onset primary lateral sclerosis. *Neurology Genetics*, 2(2), e60. Doi:10.1212/nxg.0000000000000060
- Yasojima, K., Schwab, C., McGeer, E. G., & McGeer, P. L. (1999). Up-Regulated Production and Activation of the Complement System in Alzheimer's Disease Brain. *The American Journal of Pathology*, 154(3), 927–936. Doi:10.1016/s0002-9440(10)65340-0
- Yescas, P., Huertas-Vazquez, A., Villarreal-Molina, M. T., Rasmussen, A., Tusié-Luna, M. T., López, M., *et al.* (2006). Founder effect for the Ala431Glu mutation of the presenilin 1 gene causing early-onset Alzheimer's disease in Mexican families. *Neurogenetics*, 7(3), 195–200. doi:10.1007/s10048-006-0043-3

- Yuan, A., Sasaki, T., Kumar, A., Peterhoff, C. M., Rao, M. V., Liem, R. K., *et al.* (2012). Peripherin Is a Subunit of Peripheral Nerve Neurofilaments: Implications for Differential Vulnerability of CNS and Peripheral Nervous System Axons. *Journal of Neuroscience*, 32(25), 8501–8508. Doi:10.1523/jneurosci.1081-12.2012
- Zago, S., Poletti, B., Morelli, C., Doretti, A., & Silani, V. (2011). Amyotrophic lateral sclerosis and frontotemporal dementia (ALS-FTD). *Arch Ital Biol*. 149(1), 39-56. Doi:10.4449/aib.v149i1.1263
- Zarei, S., Carr, K., Reiley, L., Diaz, K., Guerra, O., Altamirano, P. F., Pagani, W., Lodin, D., Orozco, G., & China, A. A. (2015). Comprehensive review of amyotrophic lateral sclerosis. *Surg Neurol Int*, 6, 171
- Zhang, J., Chen, C., Hua, S., Liao, H., Wang, M., Xiong, Y., & Cao, F. (2017). An updated meta-analysis of cohort studies: Diabetes and risk of Alzheimer's disease. *Diabetes Research and Clinical Practice*, 124, 41–47. Doi:10.1016/j.diabres.2016.10.024
- Zhang, Q., Gao, X., Li, C., Feliciano, C., Wang, D., Zhou, D., *et al.* (2016). Impaired Dendritic Development and Memory in Sorbs2 Knock-Out Mice. *Journal of Neuroscience*, 36(7), 2247–2260. Doi:10.1523/jneurosci.2528-15.2016
- Zhao, W., Beers, D. R., & Appel, S. H. (2013). Immune-mediated Mechanisms in the Pathoprogession of Amyotrophic Lateral Sclerosis. *Journal of Neuroimmune Pharmacology*, 8(4), 888–899. Doi:10.1007/s11481-013-9489-x
- Zhou, R., Yang, G., & Shi, Y. (2017). Dominant negative effect of the loss-of-function γ -secretase mutants on the wild-type enzyme through heterooligomerization. *Proceedings of the National Academy of Sciences*, 114(48), 12731–12736. doi:10.1073/pnas.1713605114
- Zhu, X.-C., Yu, J.-T., Jiang, T., Wang, P., Cao, L., & Tan, L. (2014). *CRI* in Alzheimer's Disease. *Molecular Neurobiology*, 51(2), 753–765. Doi:10.1007/s12035-014-8723-8
- Zhu, X. C., Tan, L., Wang, H. F., Jiang, T., Cao, L., Wang, C., Wang, J., Tan, C., Meng, X. F., & Yu, J. T. (2015). Rate of early onset Alzheimer's disease: A systematic review and meta-analysis. *Ann. Transl. Med.* 3.

Appendix I License agreements

1. Licenses from Elsevier for figures 1.5, 1.6, 1.13 and 1.14.

ELSEVIER LICENSE

TERMS AND CONDITIONS Aug 24, 2020

This Agreement between Instituto de Biomedicina de Valencia -- Jacek Szymanski ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number 4895361208560

License date Aug 24, 2020

Licensed Content Publisher Elsevier

Licensed Content Publication Neuron

Licensed Content Title The molecular pathology of Alzheimer's disease

Licensed Content Author Dennis J. Selkoe

Licensed Content Date Apr 1, 1991

Licensed Content Volume 6

Licensed Content Issue 4

Licensed Content Pages 12

Start Page 487

End Page 498

Type of Use reuse in a thesis/dissertation

Portionfigures/tables/illustrations

Number of figures/tables/illustrations 1

Formatboth print and electronic

Are you the author of this Elsevier article? No

Will you be translating? No

Title Molecular mechanisms in neurological diseases

Institution name Instituto de Biomedicina de Valencia

Expected presentation date Dec 2020

Portions Figure 1.

Instituto de Biomedicina de Valencia Jaime Roig 11

Requestor Location

Valencia, 46010 Spain

Attn: Instituto de Biomedicina de Valencia

Publisher Tax ID GB 494 6272 12

Total 0.00 EUR

Terms and Conditions

ELSEVIER LICENSE

TERMS AND CONDITIONS Aug 19, 2020

This Agreement between Instituto de Biomedicina de Valencia -- Jacek Szymanski ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number 4892470742946

License date Aug 19, 2020

Licensed Content Publisher Elsevier

Licensed Content Publication Elsevier Books

Licensed Content Title From Molecules to Networks

Licensed Content Author Monica Gireud,Natalie Sirisaengtaksin,Andrew J.

Bean

Licensed Content Date Jan 1, 2014

Licensed Content Pages 23

Start Page 639

End Page 661

Type of Use reuse in a thesis/dissertation

Portionfigures/tables/illustrations

Number of figures/tables/illustrations 1

Formatboth print and electronic

Are you the author of this Elsevier chapter? No

Will you be translating? No

Title Molecular mechanisms in neurological diseases

Institution name Instituto de Biomedicina de Valencia

Expected presentation date Dec 2020

Portions Figure 21.1
Instituto de Biomedicina de Valencia Jaime Roig 11
Requestor Location
Valencia, 46010 Spain
Attn: Instituto de Biomedicina de Valencia
Publisher Tax ID GB 494 6272 12
Total 0.00 EUR
Terms and Conditions

ELSEVIER LICENSE

TERMS AND CONDITIONS

Aug 27, 2020

This Agreement between Instituto de Biomedicina de Valencia -- Jacek Szymanski ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number 4897011403851

License date Aug 27, 2020

Licensed Content Publisher Elsevier

Licensed Content Publication Trends in Neurosciences

Licensed Content Title PRRT2: from Paroxysmal Disorders to Regulation of Synaptic Function

Licensed Content Author Flavia Valtorta,Fabio Benfenati,Federico Zara,Jacopo Meldolesi

Licensed Content Date Oct 1, 2016

Licensed Content Volume 39

Licensed Content Issue 10

Licensed Content Pages 12

Start Page 668

End Page 679

Type of Use reuse in a thesis/dissertation

Portion figures/tables/illustrations

Number of figures/tables/illustrations 2

Format both print and electronic

Are you the author of this Elsevier article? No

Will you be translating? No

Title Molecular mechanisms in neurological diseases

Institution name Instituto de Biomedicina de Valencia

Expected presentation date Dec 2020

Portions Figure 1 and Figure 2

Requestor Location Instituto de Biomedicina de Valencia Jaime Roig 11 Valencia, 46010 Spain

Attn: Instituto de Biomedicina de Valencia

Publisher Tax ID GB 494 6272 12

Total 0.00 EUR

Terms and Conditions

INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions

established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at <http://myaccount.copyright.com>).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.
5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier's permissions helpdesk [here](#)). No modifications can be made to any Lancet figures/tables and they must be reproduced in full.
6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.
7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.
8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.
9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.
10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.
11. No Transfer of License: This license is personal to you and may not be sublicensed,

assigned, or transferred by you to any other person without publisher's written permission.

12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those

established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

LIMITED LICENSE

The following terms and conditions apply only to specific license types:

15. Translation: This permission is granted for non-exclusive world English rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article.

16. Posting licensed content on any Website: The following terms and conditions apply as follows: Licensing material from an Elsevier journal: All content posted to the web site must maintain the copyright information line on the bottom of each image; A hyper-text must be included to the Homepage of the journal from which you are licensing at

<http://www.sciencedirect.com/science/journal/xxxxx> or the Elsevier homepage for books at

<http://www.elsevier.com>; Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

Licensing material from an Elsevier book: A hyper-text link must be included to the Elsevier homepage at <http://www.elsevier.com> . All content posted to the web site must maintain the copyright information line on the bottom of each image.

Posting licensed content on Electronic reserve: In addition to the above the following clauses are applicable: The web site must be password-protected and made available only to bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.

17. For journal authors: the following clauses are applicable in addition to the above:

Preprints:

A preprint is an author's own write-up of research results and analysis, it has not been peer-reviewed, nor has it had any other value added to it by a publisher (such as formatting, copyright, technical enhancement etc.).

Authors can share their preprints anywhere at any time. Preprints should not be added to or enhanced in any way in order to appear more like, or to substitute for, the final versions of articles however authors can update their preprints on arXiv or RePEc with their Accepted Author Manuscript (see below).

If accepted for publication, we encourage authors to link from the preprint to their formal publication via its DOI. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help users to find, access, cite and use the best available version. Please note that Cell Press, The Lancet and some society-owned have different preprint policies. Information on these policies is available on the journal homepage.

Accepted Author Manuscripts: An accepted author manuscript is the manuscript of an article that has been accepted for publication and which typically includes author-incorporated changes suggested during submission, peer review and editor-author communications.

Authors can share their accepted author manuscript: immediately via their non-commercial person homepage or blog

by updating a preprint in arXiv or RePEc with the accepted manuscript

via their research institute or institutional repository for internal institutional uses or as part of an invitation-only research collaboration work-group directly by providing copies to their students or to research collaborators for their personal use

for private scholarly sharing as part of an invitation-only work group on commercial sites with which Elsevier has an agreement

After the embargo period

via non-commercial hosting platforms such as their institutional repository via commercial sites with which Elsevier has an agreement

In all cases accepted manuscripts should:

link to the formal publication via its DOI

bear a CC-BY-NC-ND license - this is easy to do

if aggregated with other manuscripts, for example in a repository or other site, be shared in alignment with our hosting policy not be added to or enhanced in any way to appear more like, or to substitute for, the published journal article.

Published journal article (JPA): A published journal article (PJA) is the definitive final record of published research that appears or will appear in the journal and embodies all value-adding publishing activities including peer review co-ordination, copy-editing, formatting, (if relevant) pagination and online enrichment.

Policies for sharing publishing journal articles differ for subscription and gold open access articles:

Subscription Articles: If you are an author, please share a link to your article rather than the full-text. Millions of researchers have access to the formal publications on ScienceDirect,

and so links will help your users to find, access, cite, and use the best available version.

Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

If you are affiliated with a library that subscribes to ScienceDirect you have additional private sharing rights for others' research accessed under that agreement. This includes use for classroom teaching and internal training at the institution (including use in course packs and courseware programs), and inclusion of the article for grant funding purposes.

Gold Open Access Articles: May be shared according to the author-selected end-user license and should contain a CrossMark logo, the end user license, and a DOI link to the formal publication on ScienceDirect.

Please refer to Elsevier's posting policy for further information.

18. For book authors the following clauses are applicable in addition to the above:

Authors are permitted to place a brief summary of their work online only. You are not

allowed to download and post the published electronic version of your chapter, nor may you

scan the printed edition to create an electronic version. Posting to a repository: Authors are permitted to post a summary of their chapter only in their institution's repository.

19. Thesis/Dissertation: If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for Proquest/UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

Elsevier Open Access Terms and Conditions

You can publish open access with Elsevier in hundreds of open access journals or in nearly 2000 established subscription journals that support open access publishing. Permitted third party re-use of these open access articles is defined by the author's choice of Creative Commons user license. See our open access license policy for more information.

Terms & Conditions applicable to all Open Access articles published with Elsevier:

Any reuse of the article must not represent the author as endorsing the adaptation of the article nor should the article be modified in such a way as to damage the author's honour or reputation. If any changes have been made, such changes must be clearly indicated.

The author(s) must be appropriately credited and we ask that you include the end user license and a DOI link to the formal publication on ScienceDirect.

If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source it is the responsibility of the user to ensure their reuse complies with the terms and conditions determined by the rights holder.

Additional Terms & Conditions applicable to each Creative Commons user license:

CC BY: The CC-BY license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article and to make commercial use of the Article (including reuse and/or resale of the Article by commercial entities), provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. The full details of the license are

available at <http://creativecommons.org/licenses/by/4.0>.

CC BY NC SA: The CC BY-NC-SA license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article, provided this is not done for commercial purposes, and that the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if

changes were made and the licensor is not represented as endorsing the use made of the work. Further, any new works must be made available on the same conditions. The full details of the license are available at <http://creativecommons.org/licenses/by-nc-sa/4.0>.

CC BY NC ND: The CC BY-NC-ND license allows users to copy and distribute the Article, provided this is not done for commercial purposes and further does not permit distribution of the Article if it is changed or edited in any way, and provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, and that the licensor is not represented as endorsing the use made of the work. The

full details of the license are available at <http://creativecommons.org/licenses/by-nc-nd/4.0>.

Any commercial reuse of Open Access articles published with a CC BY NC SA or CC BY NC ND license requires permission from Elsevier and will be subject to a fee.

Commercial reuse includes:

Associating advertising with the full text of the Article Charging fees for document delivery or access

Article aggregation

Systematic distribution via e-mail lists or share buttons

Posting or linking by commercial companies for use by customers of those companies.

20. Other Conditions:

v1.10

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.


2. License agreement for figure 1.7.


RE: Requesting Permission

Data: 24.08.2020 (17:44:19 CEST)

Od: PNAS Permissions

Do: 'Jacek Szymanski'

 Pokaż źródło

 Zapisz

Thank you for your message. Permission is granted for your use of the material as described in your request. Please include a complete citation for the original PNAS article when reusing the material, and include "Copyright (2002) National Academy of Sciences, U.S.A." as a copyright note. Because this material published between 1993 and 2008, a copyright note is needed. There is no charge for this material, either. Let us know if you have any questions.

Sincerely,

Kay McLaughlin for
Diane Sullenberger
PNAS Executive Editor

-----Original Message-----

From: Jacek Szymanski <jszymanski@ibv.csic.es>
Sent: Monday, August 24, 2020 10:51 AM
To: PNAS Permissions <PNASPermissions@nas.edu>
Subject: Requesting Permission

To whom it may concern,

I wish to request permission to use a figure in my doctoral thesis.
Here is the required information:

1. Jacek Szymanski, Instituto de Biomedicina de Valencia, PhD student
2. C/Jaime Roig 11, 46010, Valencia, Spain, +34963391760, jszymanski@ibv.csic.es
3. PNAS volume number, issue number, and issue date: PNAS April 2, 2002 99 (7)
4. PNAS article title: Imaging the progression of Alzheimer pathology through the brain
5. PNAS authors' names: A. David Smith
6. Page numbers of items to be reprinted: 4135-4137
7. Figure/table number or portion of text to be reprinted: Figure 3.

Appendix II Expression vectors

1. Expression vectors used in this study.

Created with SnapGene®

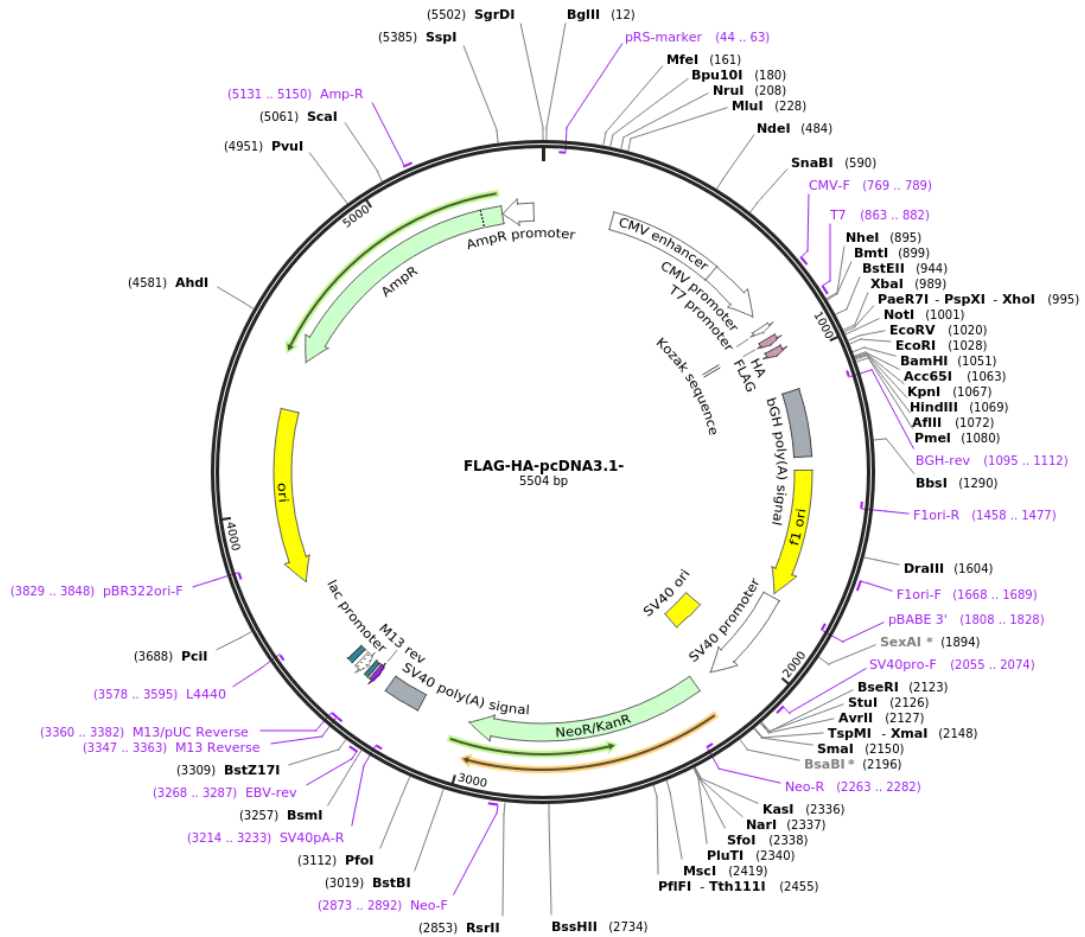


Figure A1. Vector map for expression vector FLAG-HA-pcDNA3.1 (Addgene, USA).

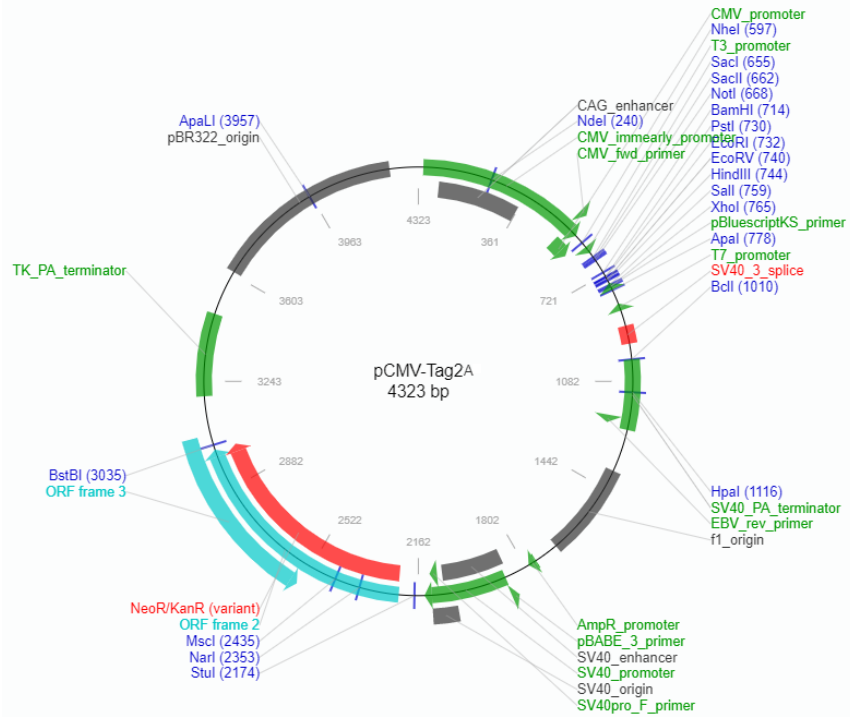


Figure A2. Vector map for expression vector pCMV-Tag2A (Donation from Jeronimo Bravo laboratory).

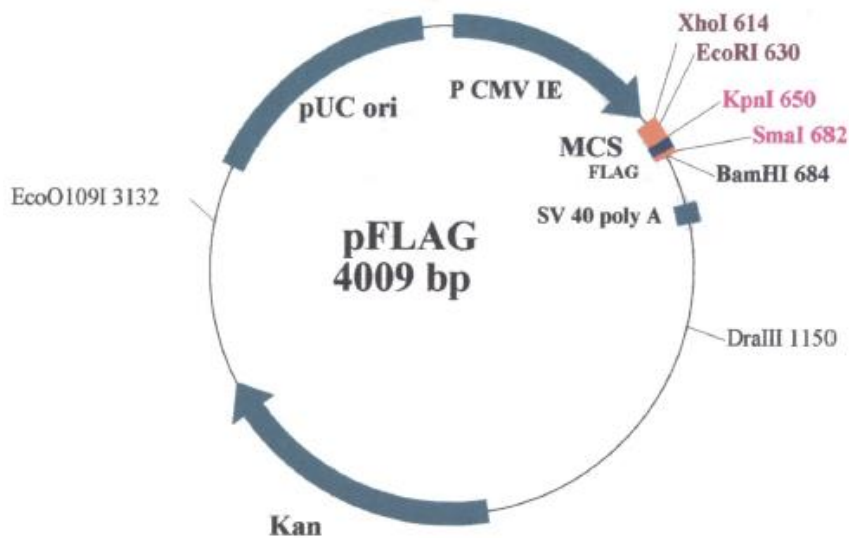


Figure A3. Vector map for expression vector pFlag – KanR; C-terminal FLAG epitope (Produced in laboratory).

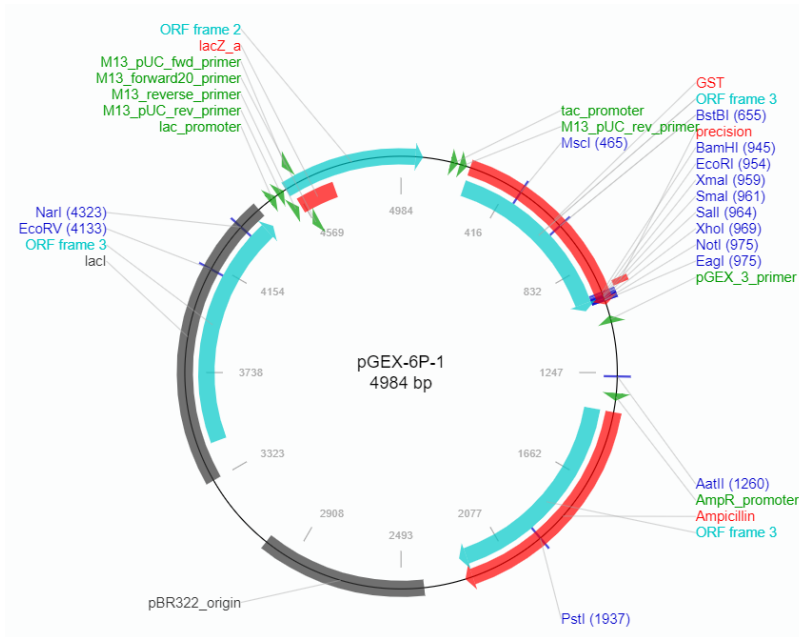


Figure A4. Vector map for expression vector pGEX-6P-1 (Addgene, USA).

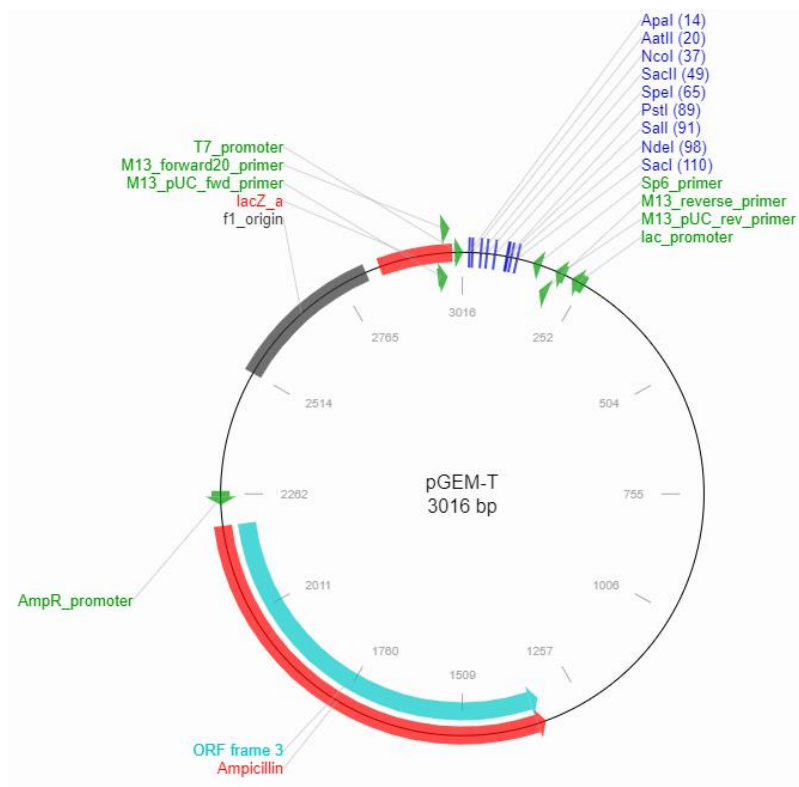


Figure A5. Vector map for expression vector pGEM-T (Sino Biological, China).

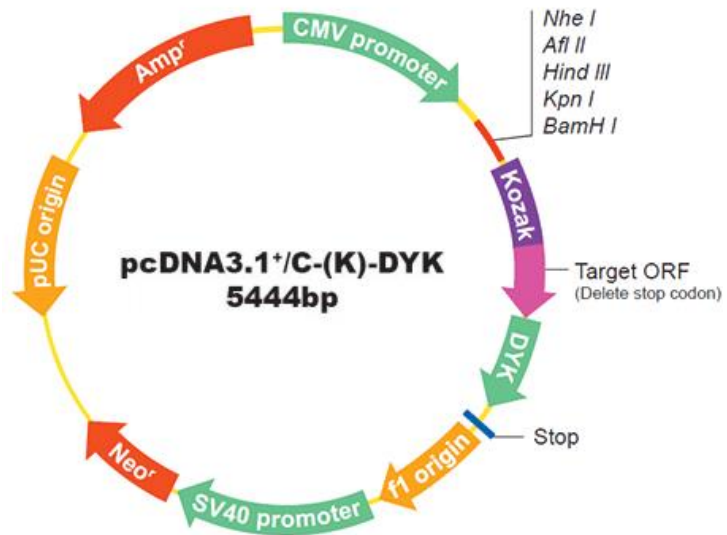


Figure A6. Vector map for expression vector pcDNA3.1+/C-(K)DYK (GenScript, USA).

Appendix III Commercial kits

Table A7. URLs to information on commercial kits used in this work.

Kit name	Producer	URL
QIAprep Spin Miniprep Kit	QIAGEN, Netherlands	https://www.qiagen.com/us/products/top-sellers/qiaprep-spin-miniprep-kit/#orderinginformation
ZymoPURE II Plasmid Midiprep	Zymo Research, USA	https://www.zymoresearch.com/collections/zymopure-plasmid-kits/products/zymopure-ii-plasmid-midiprep-kit
High Pure PCR Cleanup Micro Kit	Roche, Switzerland	https://www.sigmaaldrich.com/catalog/product/roche/hpprcuro?lang=es&region=ES
MinElute Gel Extraction Kit	QIAGEN, Netherlands	https://www.qiagen.com/dk/products/discovery-and-translational-research/dna-rna-purification/dna-purification/dna-cleanup/minelute-gel-extraction-kit/#orderinginformation
Rapid DNA Dephos & Ligation Kit	Roche, Switzerland	https://www.sigmaaldrich.com/catalog/product/roche/rdligro?lang=es&region=ES
QIAmp RNA Blood Mini Kit	QIAGEN, Netherlands	https://www.qiagen.com/us/products/discovery-and-translational-research/dna-rna-purification/rna-purification/total-rna/qiaamp-rna-blood-mini-kit/#orderinginformation
Expand Reverse Transcriptase	Roche, Switzerland	https://www.sigmaaldrich.com/catalog/product/roche/ertro?lang=es&region=ES

Appendix IV Oligonucleotide primers

Table A1. Primers used for PCR for Sanger sequencing.

Primer number	Gene	Forward primer (5'-3')	Reverse primer (5'-3')
100f/r	<i>ADAM33</i>	TGTGGCCGATCTCAT GGG	ACGTGGGTGCCTCT GACC
101f/r	<i>ADPRH</i>	GTCATTACTCTCGAC CCTCCCAGG	CTGAGCTTAAGGTG CCCAAGGG
102f/r	<i>ADPRHL1</i>	TGCTTCAGAAGAAGG GTGCT	ACGGCTCTCTTGTC AGACCT
103f/r	<i>ARHGEF10L</i>	TCAGTTTCTCTTCCTC TGGGG	CTCCTCTCACTTGTC GGGTT
104f/r	<i>AUTS2</i>	AGGGGAAGGAGGTA GGATTCT	TGATGTCAGTGGTG GTGGAG
105f/r	<i>CACNA1B</i>	GGTCAGACCCTCACG ATAGC	AAATACAGCACTGG CACCC
106f/r	<i>CRI</i>	CCTTTGATGAGTGTA AAAAGTCC	AATGTGTCGGTAAA AGAAATAGGA
107f/r	<i>DLGAP2</i>	GCAGCACTTCAATGA GGAGC	CGGCATCATCACGC ACTC
108f/r	<i>DNAJC9</i>	AAAGATCAGAAGGC GCCAAG	CTGCTGGACCTTTG CGAG
109f/r	<i>KCNN3</i>	GAAAGCGGTGGGAG AGGAG	CTCTGGGGATGAGC AGCAG
110f/r	<i>KRT9</i>	GCTCCATCTCAAACC TGCAG	TGATATCATGGCTG TGTCCTTT
111f/r	<i>MED25</i>	CTCCTTACTAGTTCCC CTCAGG	CTCACGCCCTTCT GCAG
112f/r	<i>NOTCH1</i>	GATGCATGCGTCGTT GAG	TACCGAGGATGTGG ACGAGT

Table A1. Primers used for PCR for Sanger sequencing (continued).

113f/r	<i>PARBP</i>	TCATAGGTGGCCACG TCTTT	GTTTCTTCGCCTCAC CGTTT
114f/r	<i>PRND</i>	GCTGCTCTGAGGCTA ATGTG	TGAGCCAGATCAAA GCCAGA
115f/r	<i>PRRT2</i>	GGCCACAGACCTCAG CTTAA	GGTAGGGAGCTCTG GTTGAA
116f/r	<i>PSEN1</i>	ATTAGCCTAGCGTGG TGG	GGCTTAGAATTAAC TGTAGTCTTAAGT
117f/r	<i>RYR1</i>	CCCCACACCATGTCT TCTCT	AAGAAAGGGCATGT GAAGGG
118f/r	<i>SIN3A</i>	TGGTGAACATCTCTC TCAGTGA	TGATTGTAGTCACC TGTCGGA
119f/r	<i>SIRT1</i>	AGAGGAGGCGAGGG AGGA	CCCATTTGTCTCCTTC CCCAG
120f/r	<i>SORBS2</i>	CTTGTACATCTTCAG GAACCCC	ATCAAATCCCGGAG CTGTGA
121f/r	<i>ZPR1</i>	AGATAGGGCACACA GAGCAG	TGCAGTGACTCTCT CTTCTCC

f/r – forward and reverse.

Table A2. Primers used for allele-specific PCR (ASPCR).

Primer number	Gene	Forward primer (5'-3')	Reverse primer (5'-3')	Alternative primer with 3' mismatch (5'-3')
200f/r	<i>ADPRH</i>	CCTCGAAGGAGC TTGCCCATCG	TGAGCTTAAG GTGCCCAAGG	CCTCGAAGGAG CTTGCCCATCC
201f/r	<i>CRI</i>	CAGCTTCCCTGCA CTCAAGA	GACGAGGAAC CAATGAGTCG	GACGAGGAACC AATGAGTCA
202f/r	<i>PRRT2</i>	GCCCCGAAGACC CATGCC	CTGTGTGCCCT TCTCATTCTG	GCCCCGAAGAC CCATGCT

Mismatch site indicated in bold with grey background; f/r – forward and reverse.

Table A3. Primers used for site-directed mutagenesis.

Primer number	Plasmid	Forward primer (5'-3')	Reverse primer (5'-3')	Description
300f/r	<i>ADPRH</i> (BC063883.1) in pGEM-T	TGGAGAGTTA GTG CCG CCAC AGTGATGCACT TGGCC	GGCCAAGTGC ATCACTGT GGC G GC ACTAACT CTCCA	Creation of sequence encoding ADPRH ^{D55A/D56A} variant.
301f/r	<i>ADPRH</i> (BC063883.1) in pGEM-T	GGAGCTTGCCC ACT T GAGCCTTT TTCCATGG	CCATGGAAAA AGGCTC A GTG GGCAAGCTCC	Creation of sequence encoding ADPRH ^{R295*} variant.
302f/r	<i>ADPRH</i> (BC063883.1) in pGEM-T	GGAGCTTGCCC ACC C AGCCTTT TTCCATGG	CCATGGAAAA AGGCTG G GTG GGCAAGCTCC	Creation of sequence encoding ADPRH ^{R295P} variant.
303f/r	<i>ADPRH</i> (BC063883.1) in pGEM-T	GGAGCTTGCCC ACC A AGCCTTT TTCCATGG	CCATGGAAAA AGGCT T GGTG GGCAAGCTCC	Creation of sequence encoding ADPRH ^{R295Q} variant.
304f/r	<i>CR1</i> (NM_000651.4) in pcDNA3.1 ⁺ /C- (K)DYK	TGTACTAAAG GATACGATTAC AAGGATGACG ACGATAAGTG ACTCATTGGTT CCTCG	CGAGGAACCA ATGAGTCACTT ATCGTCGTCAT CCTTGTAATCG TATCCTTTAGT ACA	Creation of sequence encoding CR1 ^{R136*} variant followed by flag epitope.

Base change indicated in bold with grey background; f/r – forward and reverse.

Table A4. Primers used for subcloning.

Primer number	Forward primer (5'-3')	Reverse primer (5'-3')	Description
400f/r	CACAG/AGC TCATGGGCG CCAGCAGCC CTA	CACAGGTA C/CGTCAGG GCGTCGTGG GTTCTG	Subcloning <i>CRI</i> (NM_000651.4) from pcDNA3.1 ⁺ /C-(K)DYK to pcDNA3.1 ⁺ /C-(K)DYK (<i>SacI</i> and <i>KpnI</i> restriction enzymes cut sites denoted by “/”)
401f/r	CACAG/AGC TCATGGGCG CCAGCAGCC CTA	CACAGGTA C/CGTG TAG CCCTTGGTG CAGCTGTA	Subcloning CR1 ^{R136*} variant from pcDNA3.1 ⁺ /C-(K)DYK to pcDNA3.1 ⁺ /C-(K)DYK (<i>SacI</i> and <i>KpnI</i> restriction enzymes cut sites denoted by “/”)
402f/r	ACTAC/TCG AGATGGAG AAGTATGTG GCTG	ATATTA/AG CTTAAGGG AAATTACAG TGTC	Subcloning <i>ADPRH</i> (BC063883.1) from pGEM-T to pFlag (<i>XhoI</i> and <i>HindIII</i> restriction enzymes cut sites denoted by “/”)
403f/r	ACTAG/GAT CCATGGAG AAGTATGTG GCTG	ATATC/TCG AGCTAAAG GGAAATTAC AGTG	Subcloning <i>ADPRH</i> (BC063883.1) from pGEM-T to pGEX-6P-1 (<i>BamHI</i> and <i>XhoI</i> restriction enzymes cut sites denoted by “/”)
404f/r	CACAG/GAT CCGATGGC AGCCAGCA GCTCT	CACAG/TCG ACTGCTACT TATACACGC CTAAGTT	Subcloning <i>PRRT2</i> (NM_145239.2) from pcDNA3.1 ⁺ /C-(K)DYK to pCMV-Tag2A (<i>BamHI</i> and <i>SalI</i> restriction enzymes cut sites denoted by “/”)
405f/r	CACAG/GAT CCGATGGC AGCCAGCA GCTCT	CACAG/TCG ACTGCTAGC ATGGGTCTT CGGGGCT	Subcloning PRRT2 ^{Q106*} variant from pcDNA3.1 ⁺ /C-(K)DYK to pCMV-Tag2A (<i>BamHI</i> and <i>SalI</i> restriction enzymes cut sites denoted by “/”)

Table A4. Primers used for subcloning (continued).

406f/r	CACAG/GAT	CACAG/TCG	Subcloning PRRT2 ^{Q163*} variant from pcDNA3.1 ⁺ /C-(K)DYK to pCMV-Tag2A (<i>Bam</i> HI and <i>Sal</i> I restriction enzymes cut sites denoted by “/”)
	CCGATGGC	ACTGCTAGG	
	AGCCAGCA	TAGGGAGCT	
	GCTCT	CTGGTTG	
407f/r	CACAG/GAT	CACAG/TCG	Subcloning PRRT2 ^{Q250*} variant from pcDNA3.1 ⁺ /C-(K)DYK to pCMV-Tag2A (<i>Bam</i> HI and <i>Sal</i> I restriction enzymes cut sites denoted by “/”)
	CCGATGGC	ACTGCTAGG	
	AGCCAGCA	AGCTGGGGT	
	GCTCT	GGCGGCT	

All restriction enzymes in table A4 are FastDigest produced by Thermo Fisher Scientific, USA; introduced restriction site indicated in bold with grey background; f/r – forward and reverse.

Table A5. Primers used for quantitative PCR.

Primer number	Target	Forward primer (5'-3')	Reverse primer (5'-3')
500f/r	<i>ADPRH</i>	CCTTCCCTGAGTCTTTTCG GT	GCAGCAAGAACAGCATCG TA
501f/r	<i>CR1</i>	GCCAGGCCTACCAACCT AAC	GCCATTCACAGGATCTGG AG
502f/r	Neomycin	CGCGGCTATCGTGGCTG GC	CGCTTGGTGGTTCGAATGG GCAGG
503f/r	<i>PRRT2</i>	GGCCATTCTGAAGCTGA AACT	CGGGGTCTCTGTGGTTTCT

f/r – forward and reverse.

Table A6. Primers used for polymorphism genotyping.

Primer number	Gene	Polymorphism examined	Forward primer (5'-3')	Reverse primer (5'-3')
600f/r	<i>APOE</i>	Allele	GGCACGGCTGTCCA AGGA	CGGGCCCCGGCCTGG TACAC
601f/r	<i>CRI</i>	rs3818361	TTCTGCTTGGTTTCC TTAGCTT	CTGCCCAGGTGAAAA GTTGT
602f/r	<i>CRI</i>	<i>CRI</i> *2 isoform	AATGTGTTTTGATTT CCCAAGATCAG	CTCAACCTCCCAAAG GTGCTA

f/r – forward and reverse.