Pathway-based genetic analyses to detect novel genes associated with frontotemporal lobar degeneration

Stéphanie Philtjens

promotor :
Prof. dr. Pieter STINISSEN,
prof. dr. M. CRUTS

co-promotor :
dr. I. GIJSELINCK

Universiteit Maastricht

Eindverhandeling voorgedragen tot het bekomen van de graad master in de biomedische wetenschappen klinische moleculaire wetenschappen

universiteit hasselt

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

FTLD	Frontotemporal lobar degeneration
FTLD-τ	Tau – positive FTLD
FTLD-TDP	TDP-43 positive FTLD
FTLD-U	Ubiquitin-positive FTLD
GRN	Granulin
VCP	Valosin-containing protein
MAPT	Microtubule-associated protein tau
CHMP2B	Chromatin modifying protein 2B
AD	Alzheimer's disease
Aβ	β-amyloid
FTD	Frontotemporal dementia
SD	Semantic dementia
PPA	Primary progressive aphasia
bv-FTD	Behavioural variant FTD
TDP-43	TAR-DNA-binding protein 43
DLDH	Dementia lacking distinctive histopathology
CBD	Corticobasal degeneration
PSP	Progressive supranuclear palsy
FTLD-MND	FTLD with motor neuron disease
ALS	Amyotrophic lateral sclerosis
FTDP-17	Frontotemporal dementia with parkinsonism linked to chromosome 17
NMD	Nonsense mediated mRNA decay
AAA-ATPase	ATPase associated with diverse cellular activities
IBMPFD	FTLD associated with inclusion myopathy and Paget disease of the bone
FTD-3	FTD linked to chromosome 3
ESCRT-III	Endosomal sorting complex required for transport III
VEGF	Vascular endothelial growth factor
APOEε4	Apolipoprotein E ε4
nNOS	Neuronal nitric oxide synthase
IPA	Ingenuity pathway analysis
CEPH	Centre d'Etude du Polymorphisme Humain
DMSO	Dimethylsulfoxide
PXN	Paxillin
CCNC	Cyclin C
NUMA1	Nuclear mitotic apparatus protein 1

PCMT1	Protein-L-isoaspartate (D-aspartate) O-methyltransferase
ATP2A2	ATPase
TIAM1	T-cell lymphoma invasion and metastasis 1
BCL2	B-cell CLL/lymphoma 2
MED1	Mediator complex subunit 1
SNP	Single nucleotide polymorphism
3’ UTR	3’ untranslated region
rs	Reference SNP

PREFACE

This thesis is the result of my senior practical training performed in the Neurodegenerative Brain Diseases Group, VIB Department of Molecular Genetics at the University of Antwerp (Belgium). For me, it was very important to perform this 30 weeks of internship on a subject that matters to me, but also in a place which could offer me the best guidance to become a great scientist. During this internship, I realized that research is not something you do on your own. It is a process in which more people work together to come to the best possible result. Therefore, I would like thank all the people who helped me and guided me during this internship.

First of all, I would like to thank Prof. dr. Christine van Broeckhoven for giving me the opportunity to conduct my internship in her research institute and for being my inspiration when I needed one. I would also like to thank my promotor Prof. dr. Marc Cruts for guiding me when I needed guidance. When I had questions you were always ready to answer them and to teach me how to be a good scientist. Thank you to my co-promotor, Ilse Gijselinck, for your trust and guidance. Githa Maes, thank you for helping me with issues in the lab and for learning me how to sequence and analyze.

I would also like to thank my intern promotor, Prof. dr. Piet Stinissen for guiding me through this internship and for being involved. Thank you to Prof. dr. Luc Michiels for critically reading my thesis and asking important questions during my thesis defense.

I would like to say a special thanks to everyone in the lab for sharing some special times with me. Especially the students, Nathalie Eysackers, Jan De Ren, Carola Sales Carbonell and Quinten Verelst.

And last but not least, I want to thank my parents and friends in Hasselt for the support they gave me. Thank you for believing in me and giving me this opportunity.

Thank you all for your support!

Stéphanie

ABSTRACT

Frontotemporal lobar degeneration (FTLD) represents a group of human diseases primarily characterized by behavioural and language disturbances that are caused by the neuronal degeneration in the frontal and temporal areas of the brain. Neuropathologically, three major distinct types of FTLD are defined: FTLD characterized by pathological tau protein deposits (FTLD-τ), FTLD characterized by TDP-43 protein deposits (FTLD-TDP), and FTLD with ubiquitin-positive neuronal inclusions of unknown nature (FTLD-U). Genetic components play a pivotal role in the aetiology of FTLD. The most common type is FTLD-TDP with mutations in the *GRN* gene, the *VCP* gene and other unidentified genes.

30-50% of FTLD patients have a positive family history of dementia. However, only 50% of these inherited cases are due to mutations in *MAPT*, *GRN*, *CHMP2B* or *VCP*. It is also not clearly understood how mutations in these genes cause FTLD or how the related disease pathways function. Importantly, for about 50% of familial FTLD cases, the genetic cause is still unknown. As a result, we looked for a plausible pathway in which FTLD-TDP genes are active and we screened other genes in this pathway for their genetic contribution to the etiology of FTLD. We were able to narrow our selection of genes from 8 genes to 3 possibly causal genes. However, it still needs to be elucidated if these genes are associated with a higher risk for developing FTLD.

SAMENVATTING

Frontotemporale kwab degeneratie (FTLD) is een algemene term voor aandoeningen die gekenmerkt worden door gedrags- en taalstoornissen. FTLD wordt veroorzaakt door neuronale degeneratie in de frontale en temporale kwabben van de hersenen. Op neuropathologisch gebied kan men FTLD indelen in drie verschillende vormen. De eerste en meest voorkomende vorm is FTLD-TDP. FTLD-TDP wordt gekarakteriseerd door deposities van het TDP-43 eiwit die tevens positief zijn voor ubiquitine. Een tweede vorm is FTLD-τ en wordt gekenmerkt door tau positieve eiwit deposities. De derde en laatste vorm van FTLD is FTLD-U. Deze vorm is voornamelijk gekarakteriseerd door ubiquitine-positieve neuronale inclusies waarvan de samenstelling nog onbekend is.

Ongeveer de helft van alle FTLD patiënten vertonen een positieve familiale geschiedenis voor dementie. Maar slechts de helft van deze personen hebben een mutatie in één van de gekende FTLD genen (*MAPT, GRN, VCP* of *CHMP2B*). Dit wil zeggen dat er dus zeker mutaties in andere genen aan de oorsprong liggen van het ontstaan van FTLD. Hiervoor gingen we op zoek naar een pathway waarin de FTLD-TDP genen *GRN* en *TARDBP* actief zijn. De genen die zich ook in de pathways bevonden werden gescreend op genomisch DNA van 23 familiale FTLD patiënten. Door deze mutatieanalyse waren we in staat om drie genen te selecteren die mogelijk betrokken zijn in het ontstaan van FTLD. De andere vijf geselecteerde genen kunnen risico varianten zijn. Maar dit aspect moet in de nabije toekomst nog bewezen worden.

1 INTRODUCTION

Frontotemporal lobar degeneration (FTLD) represents a group of human diseases primarily characterized by behavioural and language disturbances that are caused by the neuronal degeneration in the frontal and temporal areas of the brain. Neuropathologically, three major distinct types of FTLD are defined: FTLD characterized by pathological tau protein deposits (FTLD-τ), FTLD characterized by TDP-43 protein deposits (FTLD-TDP), and FTLD with ubiquitin-positive neuronal inclusions of unknown nature (FTLD-U). Genetic components play a pivotal role in the aetiology of FTLD. The most common type is FTLD-TDP with mutations in the granulin gene (*GRN*), the valosin-containing protein gene (*VCP*) and other unidentified genes. Mutations in the microtubule associated protein tau gene (*MAPT*) are associated with FTLD-τ while mutations in the chromatin modifying protein 2B gene (*CHMP2B*) are associated with FTLD-U.

The goal of this study was to look for new FTLD-TDP genes. By making use of a bioinformatic tool, we looked for putative pathways connecting *GRN* and *TARDBP*. The genes in these pathways were possible candidates for sequence analysis. In the first part of the introduction, a general overview will be given about dementia, FTLD and its pathology. Subsequently, the genetic basis of FTLD will be explained with a short introduction of the known FTLD genes, their locus and possible risk factors. Next, a closer look on the *TARDBP* gene will be given followed by an outline of the current study.

1.1 Frontotemporal lobar degeneration

In 2005, it was estimated that approximately 24 million people suffered from dementia worldwide. Due to the increasing life expectancy, it is becoming clear that this number will double every 20 years to 42 million by 2020 and 81 million by 2040 [1].

Neurodegenerative dementia is a general symptom occurring in a group of disorders that cause irreversible cognitive decline which results from damage of the cells in the brain. Before one can speak of dementia, a disorder must cause decline in some essential cognitive functions such as memory, ability to generate coherent speech or understand spoken or written language, capacity to plan and carry out complex tasks, and ability to process and interpret visual information. This decline must be severe enough to interfere with everyday life. The most common cause of dementia is Alzheimer's disease (AD) which accounts for 50-70% of all the dementia cases and is characterized by β-amyloid (Aβ) plaques and neurofibrillary tangles in brain [2].

The second most common cause of neurodegenerative dementia in people under the age of 65 years is FTLD, accounting for 5-10% of all dementia patients and 10-20% of patients younger than 65 years of age [3, 4]. Clinical characteristics of FTLD are changes in personality and behaviour, but also impairment of language and executive dysfunction. Pathologically, the disease affects the frontal and temporal lobes of the brain, generally resulting in three different clinical syndromes: frontotemporal dementia (FTD), semantic dementia (SD) and primary progressive aphasia (PPA) [5]. The most common clinical manifestation of FTLD is FTD. FTD, also termed behavioural variant FTD (bv-FTD) occurs typically between 45 and 65 years of age and is characterized by changes in personality and social conduct, with a relative preservation of language and memory [6]. This syndrome is associated with atrophy of the prefrontal and anterotemporal cortex. The other two FTLD syndromes, SD and PPA, are characterized by progressive changes in language function. This language dysfunction precedes the behavioural changes that occur in FTLD patients. SD patients lack the ability to recognize the meaning of words and visual stimuli and they are characterized by a loss of verbal and nonverbal skills. Ultimately, these patients will become mute. In SD, the presence of bilateral atrophy of the middle and infero-temporal cortex can be detected. PPA, on the other hand, is characterized by progressive word retrieval difficulties resulting in reduction of speech production. This last general FTLD syndrome is associated with asymmetric atrophy of the left frontal and temporal cortices [5, 7]. In FTLD, death often occurs after a disease duration of on average 1 to 20 years [8].

Based on the neuropathological lesions FTLD can be divided in three major groups. The first and most common FTLD form is FTLD-TDP. FTLD-TDP is characterized by tau- and α-synuclein-negative neuronal inclusions. These inclusions are immunoreactive to ubiquitin antibodies and contain aggregated TAR-DNA-binding protein-43 (TDP-43) [9, 10]. The term FTLD-U is also often used for this subset of patients. However, according to the most recent nomenclature guidelines FTLD-U patients are characterized by ubiquitin-positive neuronal inclusions without deposition of the TDP-43 protein and appears to be a different subset of FTLD [11-13]. The second most common form of FTLD is FTLD-τ. This form is characterized by neuronal and/or glial deposits of hyperphosphorylated tau proteins and is consequently referred to as a tauopathy [5]. A small percentage (<5%) of FTLD patients is classified as dementia lacking distinctive histopathology (DLDH), since no neuropathological inclusions could be detected [14].

Other syndromes belonging to the clinicopathological spectrum of FTLD are corticobasal degeneration (CBD), progressive supranuclear palsy (PSP) and FTLD with motor neuron disease (FTLD-MND) [5]. Both CBD and PSP are clinical manifestations of FTLD-τ and are characterized by cortical and striatal tau-positive neuronal and glial lesions for CBD and atrophy of brainstem and basal ganglia for PSP [15, 16]. Patients with FTLD-

MND do not only show changes in behaviour but also in motor functioning. FTLD-MND patients thus represent with characteristics of FTLD and amyotrophic lateral sclerosis (ALS). ALS is the most common form of MND and it generally includes loss of the upper en lower motor neurons [5, 17].

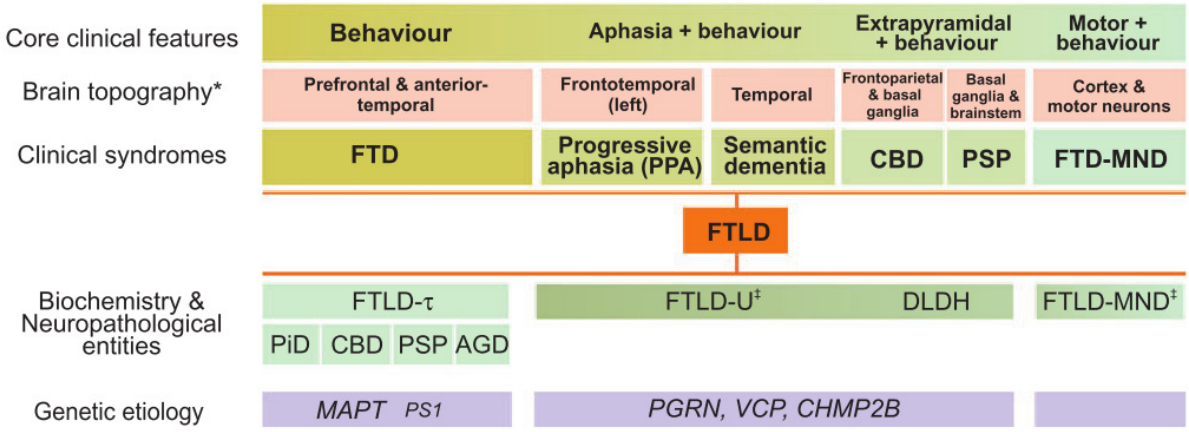


Figure 1: Overview of FTLD and his clinical syndromes, brain topography, clinical features, neuropathological characteristics and genetic etiology. AGD: argyrophilic grain disease; ALS: amyotrophic lateral sclerosis: PiD: Pick's disease; PSP: progressive supranulear palsy; CBD: corticobasal degeneration; FTD: frontotemporal dementia; FTLD: frontotemporal lobar degeneration; MND: motor neuron disease (* evident on both neuropathology and neuroimaging; ‡ the inclusion proteopathy could be TDP-43) [5].

1.2 Genetics of FTLD

After age, family history of dementia is the second most important risk factor for FTLD. About 30-50% of FTLD patients have a positive family history of dementia, most of these diseases segregate as an autosomal dominant trait throughout families [18, 19]. This suggests that genetics play a major role in the etiology of FTLD. However, only 50% of these inherited forms of FTLD are due to mutations in the known FTLD genes, i.e. microtubule-associated protein tau (*MAPT*) gene, the granulin (*GRN*) gene, the valosin-containing protein (*VCP*) gene and the charged multivesicular body protein 2B or chromatin modifying protein 2B (*CHMP2B*) gene [5]. Mutations in *MAPT* are associated with FTLD-τ while mutations in *GRN* and *VCP* are linked to FTLD-TDP. Mutations in *CHMP2B* on the other hand are associated with FTLD-UPS or FTLD-U without TDP-43 inclusions [13].

1.2.1 Causal genes

There are four causal genes for FTLD known to date. These four FTLD genes are, as described above, *GRN*, *VCP*, *CHMP2B* and *MAPT*. At this moment, there is no proof that these genes are connected through a common pathway leading to neurodegeneration. However, it is possible that *GRN* and *VCP* act in the same pathway because they both result in FTLD-TDP with tau- and α-synuclein negative, ubiguitin and TDP-43 positive deposits [20]. Mutations in the other two genes, *MAPT* and *CHMP2B*, do not result in the deposition of this TDP-43 protein and may thus be active in a completely different pathway than *GRN* and

VCP. However, both genes may not be active in the same pathway as each other since the neuropathology of *MAPT* mutations is completely different from the neuropathology of *CHMP2B* mutations. Mutations in *MAPT* results in tau positive deposits instead of ubiquitin and TDP-43 positive inclusions. The same is true for mutations in *CHMP2B*. Mutation in *CHMP2B* results in ubiquitin positive inclusion without immunoreactivity against tau or TDP-43 [12, 13]. However, it needs to be noticed that this is just a hypothesis and not a hypothesis proven by other researchers. The four FTLD genes will be explained in more detail below.

1.2.1.1 MAPT

The *MAPT* gene is located on chromosome 17q21 and FTLD with mutations in this gene were originally identified in frontotemporal dementia with parkinsonism linked to chromosome 17 (FTDP-17) [21]. Tau is a protein that is constitutively expressed in the nervous system, both central and peripheral. The protein is enriched in the axons of mature and growing neurons in the central nervous system [22, 23]. A mutation in the *MAPT* gene results in deposition of hyperphosphorylated tau in insoluble filaments in the brain. This abnormal tau protein leads to a loss of ability to assemble tubulin and to stabilise microtubules. Since the major feature in this disease is an abnormality in the tau protein, this form of FTLD is referred to as a tauopathy [24, 25]. Currently, 44 *MAPT* mutations are known in 131 families [26]. Almost every *MAPT* mutation is located in the C-terminus of the protein and include missense, silent and intronic variations. After overexpression, the tau protein reduces the attachment of kinesin to microtubules [27].

1.2.1.2 GRN

Since some FTLD families showed linkage to chromosome 17 but no tau-positive inclusions, mutations in the *GRN* gene were established. The *GRN* gene is located 1.7 Mb centromeric to the *MAPT* gene on chromosome 17q21. This gene consists of 13 exons and encodes a 593-amino-acid secreted glycoprotein of 68.5 kDa [19, 28]. *GRN* is a growth factor, important in the regulation of cell growth and cell cycle progression, and plays a role in wound healing, development, inflammation and tumorigenesis [28, 29].

Since the *GRN* gene is located on chromosome 17, FTLD with a mutation in this gene is referred to as FTLDU-17. To date, there are 64 *GRN* mutations identified in 171 families [26]. It is seen that the majority of the mutations in this gene are nonsense, frameshift and splice site mutations. These mutations result in premature stop codons and lead to degradation of mutant transcripts by nonsense-mediated mRNA decay (NMD). Briefly, NMD is a surveillance mechanism that degrades mRNA with premature termination stop codons. This mechanism prevents the production of C-truncated and nonsensical proteins. As a result of NMD, no mutant *GRN* transcript could be found and there was a reduction of 30-

50% in the levels of *GRN* protein [30-32]. Although it is established that the disease results from a reduction in functional GRN instead of the production of mutant protein, the mechanism by which *GRN* mutations lead to FTLD is not understood [14].

1.2.1.3 VCP

VCP is a gene that belongs to the AAA-ATPase gene superfamily (ATPase Associated with diverse cellular Activities) [33, 34]. *VCP* functions as a molecular chaperone in cellular processes such as ubiquitin-dependent protein degradation, stress responses and programmed cell death. This means that VCP is able to bind to the polyubiquitin chains and untheters the ubiquitinated proteins from their binding partners. This process facilitates transport to the ubiquitin proteasome system [33, 35, 36]. The *VCP* gene is located on chromosome 9p21-22. Mutations in the *VCP* gene were found in patients with inclusion body myopathy, Paget disease of the bone and frontotemporal dementia (IBMPFD). IBMPFD is a rare autosomal-dominant neurodegenerative disorder with early age of onset [20, 37]. This disorder is characterized by variable penetrance of an unusual triad of clinical features (myopathy, Paget's disease of the bone and FTD). FTD only affects 31-37% of affected individuals and is characterized by language and/or behavioural dysfunction with relative preservation of memory [38, 39]. Currently there are 12 mutations in 29 families known in the *VCP* gene [26].

1.2.1.4 CHMP2B

A mutation in the *CHMP2B* gene was first described in a Danish family which showed linkage to chromosome 3p11.2 (FTD-3) [40]. The *CHMP2B* gene is a member of the chromatin modifying protein or charged multivesicular body protein family and is a component of the endosomal sorting complex required for transport III (ESCRT-III). This complex is involved in the formation of endocytic multivesicular bodies and in the degradation of surface proteins. Nevertheless, mutations in the *CHMP2B* gene appear to be a rare cause of FTLD, only four mutations in four families are known in this gene to date [26, 41, 42].

1.2.2 Genetic risk factors

As for other neurodegenerative disorders such as AD, researchers looked for possible risk factors that increase the risk for developing FTLD. They investigated, for example, what the role was of factors such as vascular endothelial growth factor (*VEGF*), apolipoprotein E ε4 (*APOE*) and neuronal nitric oxide synthase (*nNOS*) in the development of FTLD. However, conflicting studies exist for each of these genes. There is one association that is strongly proven and that is the association between the *MAPT* H1 haplotype and PSP. A brief review of these factors will be given in the following sections.

1.2.2.1 VEGF

Vascular endothelial growth factor (*VEGF*) is a growth factor that plays an important role in the angiogenesis and vasculogenesis. *VEGF* has recently been implicated in neuronal survival, neuroprotection, regeneration, growth, differentiation and axonal outgrowth [43]. Evidence exists of an increased immunoreactivity in the cortex of patients with AD. They also found expression of *VEGF* in reactive astrocytes and endothelial cells surrounding amyloid deposits in AD brains [44]. Thus, it was seen that *VEGF* plays a role in neurodegenerative disorders. For this reason, Borroni et al investigated whether *VEGF* also plays a role in FTLD [45]. They investigated 161 subjects with FTD, 56 with CBD syndrome and 57 with PSP. They calculated the genotype and allele frequencies of four known polymorphisms in the *VEGF* promoter in these patients and in 216 age-matched healthy subjects. From this study, they were able to conclude that *VEGF* gene variability is a susceptibility factor for sporadic FTLD. However, they report that this is the case in an Italian population and that more research on this topic is needed [45].

1.2.2.2 APOE

The APOE protein is a major protein in the central nervous system and is essential for lipid transport and cholesterol homeostasis. It is repeatedly demonstrated that the *APOE* ε4 allele is consistently associated with an increased risk to develop AD [46]. There are over twenty studies performed to investigate whether the *APOE* ε4 allele frequency is increased in FTLD [47]. However, no consensus can be reached whether *APOE* ε4 results in an increased risk for FTLD. Rosso et al [48] and Short et al [49] both suggested that the *APOE* ε4 allele frequency is increased in patients with SD. However, Pickering – Brown et al [50] showed that the *APOE* ε4 allele frequency is normal in FTD and SD, but slightly increased in progressive aphasia. Srinivasan et al [47] showed that there was a significant increase in the *APOE* ε4 allele frequency in FTLD patients compared to normal controls, which suggests that the *APOE* ε4 allele acts as a risk factor for FTLD, bus probably to a lesser extent than is seen in AD [47].

1.2.2.3 nNOS

Neuronal nitric oxide synthase (nNOS) is responsible for the generation of nitric oxide inside the cells of the brain. Earlier research has proven a remarkable loss of neurons expressing nNOS in the entorhinal cortex layer II and in the hippocampus of AD patients. This suggests that neurons that express *NOS1*, the gene encoding nNOS, are highly susceptible to neurodegeneration [51]. Because of this link between nNOS and neurodegeneration, Venturelli et al [52] looked which effect the *NOS1* C276T polymorphism has on FTLD. They showed that the presence of the T allele almost doubled the risk of developing FTLD. They further state this important role for nNOS in FTLD by saying that peroxynitrite, which is

derived from nitric oxide generated by NOS, activates several kinases such as MAP-kinase pathway. This pathway is involved in tau phosphorylation and thus in the generation of NFTs [52, 53]. So they prove that nNOS increases the risk on developing FTLD.

1.2.2.4 MAPT

PSP is a subtype of FTLD which is neuropathologically characterized by deposition of the tau protein. However, the tau deposits in PSP differ from those in AD or FTLD since they are straight filaments containing only 4-repeat tau isoforms. Baker et al [54] showed that the more common *MAPT* H1 haplotype is significantly over-represented in patients with PSP. This confirmed the earlier reports of Conrad et al who demonstrated the presence of an association between *MAPT* and PSP [55]. Baker et al gave three possibilities for explaining the association between *MAPT* and PSP. The first explanation is that there may be crucial differences between the two haplotypes in terms of expression of the tau protein. A second explanation, according to Baker et al is that there are differences in the splicing of the two cognate proteins. The third and last possible explanation for this association is that a pathogenic, non-coding, mutation has occurred on the H1 background. However, which of these explanations is correct needs to be verified [54].

1.3 TDP-43

Previous research has established that the major constituent of the ubiquitin-positive inclusions in FTLD-TDP patients is TDP-43 [10]. However, it needs to be noted that this protein is not detected in patients with *CHMP2B* mutations [56]. *TARDBP*, the gene encoding the TDP-43 protein is located on chromosome 1p36. TDP-43 is a nuclear ribonucleoprotein that is involved in the regulation of transcription and alternative splicing and appears to be ubiquitinated and hyperphosphorylated in these neuronal inclusions [10, 57]. Patients with a mutation in the *GRN* gene show TDP-43 pathology in the microglial cells of the white matter [58]. Aggregation of TDP-43 is not only seen in FTLD-TDP patients but also in other neurodegenerative diseases such as AD, ALS and Huntington's disease [59]. Based on frequency and distribution of TDP-43 inclusions, four different FTLD-TDP subtypes were proposed [14] (Figure 2).

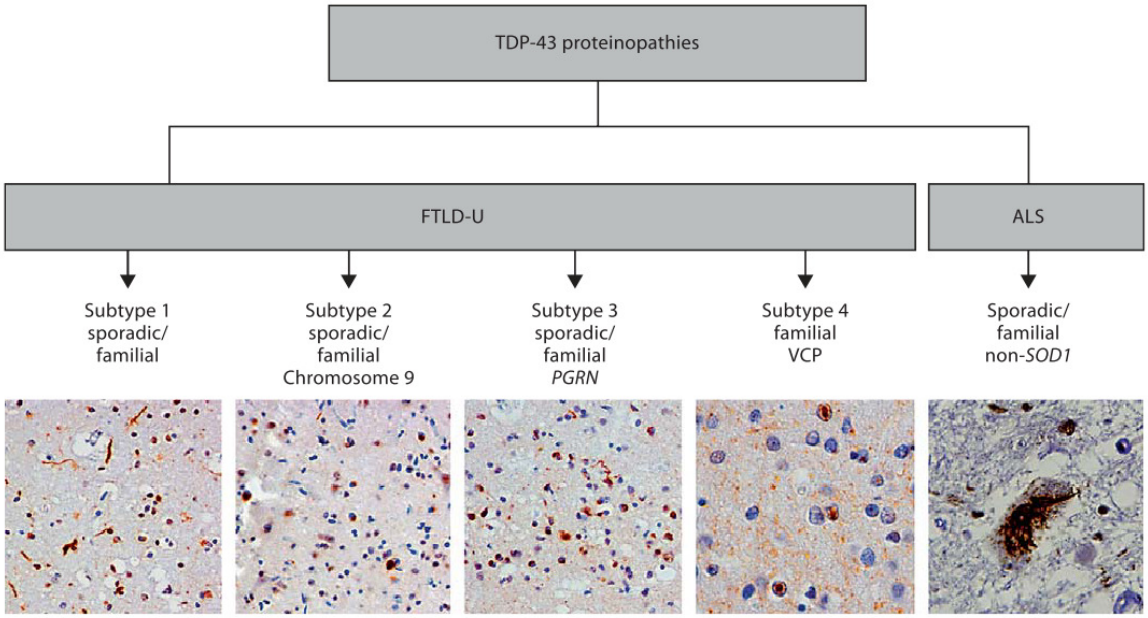


Figure 2: Classification of TDP-43 proteinopathies. Based on the TDP-43 immunohistochemical profile, four histological subtypes of FTLD-U can be described. Subtype 1 occurs both in sporadic and familial FTLD-U with no mutation in *CHMP2B*, *GRN* or *VCP*. Subtype 2 also appears in both sporadic and familial cases of FTLD-U but mutations are seen in chromosome 9. No gene is known yet for chromosome 9 in FTLD-U. Subtype 3 is seen in patients with mutations in the *GRN* (*PGRN*) gene and subtype 4 only occurs in familial FTLD-U patients with mutations in *VCP*. TDP-43 deposition also occurs in ALS. Images below each subtype and ALS illustrate histological features by anti-TDP-43 immunohistochemistry [60].

1.4 Aim of the study

About 50% of the familial FTLD cases are due to mutations in the four known FTLD genes (*GRN, CHMP2B, VCP* and *MAPT*). As a result, we claim that other genes must exist which result in FTLD when mutated. Therefore, we looked for new causal genes for FTLD in this study. First we made a selection of eight genes based on expression profiles and interactions with the known FTLD genes (*GRN*, *VCP* and *TARDBP*). For the expression profiles, we looked which genes showed a significantly increased or decreased expression in the brains of FTLD patients with a *GRN* mutation. Then we looked which of these genes have any direct or indirect interactions with *GRN*, *VCP* or *TARDBP* based on known protein-protein interactions in the literature. For these interactions, we made use of a bioinformatics tool called IPA. Next, the exons of the eight selected genes were amplified by means of PCR and subsequently sequenced. This sequencing was done to establish the presence of novel causal variants in a selection of familial patients of a Belgian FTLD patient cohort.

2 MATERIALS AND METHODS

2.1 Patient samples

Patient samples were selected from the Middelheim prospective study of neurodegenerative and vascular dementia in Belgium [61]. Inclusion criteria for this study were a primary diagnosis of FTLD, and a documented family history of dementia. Of the 170 FTLD patients in the Middelheim study, 23 were included in this research. The age at onset varied from 47 to 90 years and the disease duration varied from 1 to 12 years. Male to female ratio was 1.75 (Table 1).

Table 1: The patient samples used in this research.

Patient id	M/F	Diagnosis	Subclass	AAO	AAD	DD
d1255	M	FTLD	FTD	63	65	2
d1300	M	FTLD	FTD	90	94	4
d1520	F	FTLD	FTD	58	61	3
d1635	M	FTLD	FTD	66	68	2
d1973	F	FTLD	PNFA/FTD	77	84	7
d1976	M	FTLD	FTD	65	68	3
d2035	F	FTLD	PNFA	58	64	6
d2125	M	FTLD	pos FTD	67	71	4
d3949	M	FTLD	FTD	75	76	1
d4703	M	FTLD	FTD	65	66	1
d5772	F	FTLD	FTD	65	68	3
d5926	F	FTLD	FTD	72	75	3
d5945	M	FTLD	FTD	57	60	3
d6177	M	FTLD	pos FTD	74	74	<1
d6236	F	FTLD	FTD	74	74	<1
d6886	M	FTLD	FTD	57	69	12
d7145	M	FTLD	FTD	47	50	3
d7214	M	FTLD		79	80	1
d7267	F	FTLD	pos FTD	66	71	5
d7808	M	FTLD	pos FTD	64	68	4
d10197	M	FTLD	FTD	51		
d10321	M	FTLD	prob FTD	75	77	2
d10406	F	FTLD	PPA	63		

M/F: male/female; PNFA: Progressive non-fluent aphasia; pos FTD: possible FTD; prob FTD: probable FTD; AAO: age at onset; AAD: age at death; DD: disease duration.

These patients were negative for mutations in *GRN*, *VCP* and *CHMP2B,* with the exception of patient d3949, who has a mutation in the *GRN* gene (IVS1+189C>T). Nevertheless, it is important to notice that this mutation is no causal mutation, this means that this mutation does not lead to a change at the protein level and does not results in FTLD.

2.2 Selection of pathway and genes

The genes that resulted from co-expression and transcriptomic studies were submitted in Ingenuity Pathway Analysis (IPA) together with the *GRN*, *VCP* and *TARDBP* genes. Next, we looked which of these genes were directly or indirectly connected to one of the three target genes (*GRN*, *VCP* and *TARDBP*). Genes with an interaction were scored '+1'. Genes also received a '+1' score if they were located in one of the known FTLD-associated loci with unidentified gene defect, such as the ALSFTLD2 locus on chromosome 9p. If the genes showed a differential expression in transcriptomic studies comparing brains of FTLD patients with a *GRN* mutation and healthy controls, they also received a '+1' score. Next we looked which of these genes with the highest score were involved in pathways that started with the *GRN* gene and led to the *TARDBP* gene. These genes were selected for sequencing.

2.3 Primer design and polymerase chain reaction

The PCR primers for the selected genes were designed by making use of a bioinformatic tool Primer3. The resulting primers were optimized to find out which annealing temperature is the ideal temperature for each primer pair. To do this, we first performed a PCR using Titanium Taq DNA-polymerase with an annealing temperature of 68°C. PCR was carried out on 20 ng/µl genomic DNA of Centre d'Etude du Polymorphisme Humain (CEPH) sample 1333.01:9 in a total volume of 15 µl containing 10x Titanium-buffer (VIB, Antwerp, Belgium), 10 mM dNTP, 5U/µl Titanium Taq (50x Titanium Taq DNA polymerase, Clonetech, Westburg, Netherlands), water, 100 µM forward and 100 µM reverse primer. On the primer sets that did not work with this condition, a gradient PCR using Titanium Taq was carried out. The PCR mix was identical to that used in a standard PCR using Titanium Taq. For the primer sets that did not give the wanted result with one of these PCRs, a PCR using Platinum Taq was performed. This PCR was carried out on 20 ng/µl genomic DNA of CEPH sample 1333.01:9 in a total volume of 15 µl. The reaction mixture of this PCR existed of 10x Platinum buffer (10x PCR Rxn Buffer (-$MgCl_2$), Invitrogen, California, USA), 50 mM $MgCl_2$ (Invitrogen, California, USA), 10 mM dNTP, 10 pmol forward primer, 10 pmol reverse primer, water and 5U/µl Platinum taq (Platinum® Taq DNA polymerase, Invitrogen, California, USA). If this PCR did not work out, further optimizations were carried out by changing the concentration of $MgCl_2$ (1mM, 2mM or 3mM) or by adding dimethylsulfoxide (DMSO; 10%) or Betaïne (1M). Next, the optimized PCR reactions were carried out on genomic DNA of the FTLD patients.

To verify the size of the PCR product, the samples were separated on a 1.33% agarose gel and visualized using UV illumination.

2.4 DNA-sequencing

PCR products were purified before the sequencing reaction took place. This purification is performed to remove the overload of dNTPs and primers after the PCR reaction. The purification mix consists of TE^{-4}, ExoSAP-IT® (USB Corporation, Cleveland, Ohio, USA) and the PCR product. After the purification reaction the resultant presequencing product was diluted using water depending on the intensity of the bands on the gel. This diluted PCR product was used in the sequencing reaction. The sequencing mix consists of 5 x sequencing buffer, 10 pmol primer and big dye. The big dye is a combination of dNTPs and labeled ddNTPs. Afterwards, the sequencing reaction mixture was purified to get rid of the overload of free PCR product, primer and dNTPs. This is done by making use of 85% ethanol and magnetic beads which are carboxylate-modified polymer coated. These magnetic beads become electrostatic under the influence of ethanol and bind DNA. As a result, the DNA is pulled to the outer layer of the well. The free dNTPs are too small to bind to these beads and stay in solution. This solution will disappear and only the DNA stays in the well. The purified sample is loaded onto an ABI 3700 fragment analyzer. Here, the solution will go either into a long capillary for sequences of 800 to 900 base pairs or into a short capillary for sequences of 500 base pairs. This capillary results in a detector which detects the signal from the fluorescently labeled ddNTPs. The analyses of the sequences is done by making use of SeqMan, a program of the DNASTAR DNA analysis package.

3 RESULTS AND DISCUSSION

3.1 Gene selection through IPA

The selection of the 500 best functional candidate genes for FTLD-TDP was imported in the IPA software program. First, we investigated which paths in the interactome connected *GRN* with *TARDBP*. We identified paths composed of 2, 3 and 4 intermediate nodes. More than 4 nodes was computationally too demanding and was not investigated. Next, we looked which of these genes in the networks given by IPA were also present in the list of top score genes that we had generated earlier. The IPA analysis of two nodes between *GRN* and *TARDBP* led to a selection of eight genes which were differently expressed compared to *GRN* (over- or underexpressed) (table 2). These eight genes were *PXN*, *CCNC*, *NUMA1*, *PCMT1*, *ATP2A2*, *TIAM1*, *BCL2* and *MED1*. Table 2 shows the chromosomal location of these genes, their gene name and gene number and what their top score was based on known interactions with *GRN*, *VCP* and *TARDBP*.

Table 2: The eight genes selected after analysis with IPA.

Gene id	Ensembl nr	Gene symbol	Gene name	Chromosome	Score
NC_000012.11	ENSG00000089159	*PXN*	Paxillin	12q24.31	4
NC_000006.11	ENSG00000112237	*CCNC*	Cyclin C	6q21	3
NC_000011.9	ENSG00000137497	*NUMA1*	Nuclear mitotic apparatus protein 1	11q13	3
NC_000006.11	ENSG00000120265	*PCMT1*	Protein-L-isoaspartate (D-aspartate) O-methyltransferase	6q24-q25	3
NC_000012.11	ENSG00000174437	*ATP2A2*	ATPase (Ca++ transporting, cardiac muscle, slow twitch 2)	12q23-q24.1	3
NC_000021.8	ENSG00000156299	*TIAM1*	T-cell lymphoma invasion and metastasis 1	21q22.11	3
NC_000018.9	ENSG00000171791	*BCL2*	B-cell CLL/lymphoma 2	18q21.3	3
NC_000017.10		*MED1*	Mediator complex subunit 1	17q12-q21.1	2

Because IPA generates pathways based on known interactions in the literature, we were able to generate a possible FTLD disease pathway. In Figure 3, all the interactions between the eight selected genes and the two FTLD-TDP genes *GRN* and *TARDBP* are shown in a pathway model.

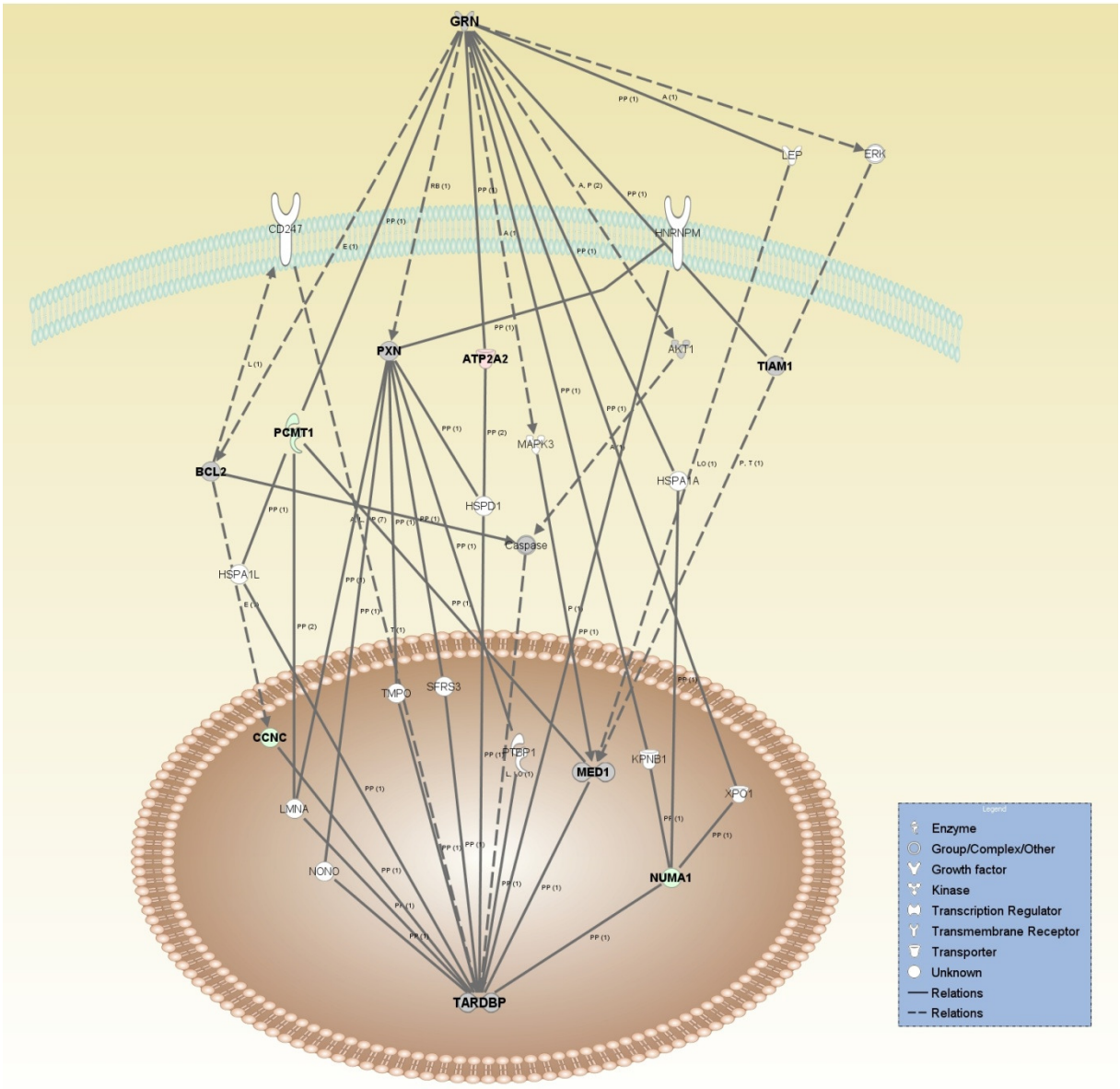


Figure 3: IPA pathway with two FTLD-TDP genes, *GRN* and *TARDBP*, and the eight genes sequenced in this study. Gene interactions which make it possible to connect *GRN* and *TARDBP* with only two genes in between. The eight genes that were selected are shown in bold (*PXN, BCL2, PCMT1, MED1, NUMA1, TIAM1, ATP2A2* and *CCNC*). Red: down-regulated (foldchange >1); green: up-regulated (foldchange >1); gray: up- or down-regulated but foldchange is <1, so the difference is small.

3.2 Mutation analysis of selected genes

The eight selected genes were sequenced to find out whether a variation or single nucleotide polymorphism (SNP) was present that could possibly cause FTLD in our patient population. For all the eight genes, we screened a series of 23 familial FTLD patients in which mutations in all known FTLD genes were excluded. Due to a lack of time, only novel variants in coding regions of exons were sequenced in 96 neurologically healthy individuals. In the next sections, the results for all eight genes are given.

3.2.1 Non-coding novel variant in *PCMT1*

PCMT1 is located on the long arm of chromosome 6 (6q24-q25) with seven coding exons and one non-coding exon in the 3' untranslated region (UTR). The encoded protein (PCMT1) is responsible for the repair of proteins that have been damaged by deamidation of asparagines or isomerization of aspartate. This deamidation results in the spontaneous generation of L-isoaspartyl residues within peptides and proteins during aging [62, 63].

We identified five SNPs, including one novel (Table 3). One SNP in exon 5 predicted a p.Val120Ile amino acid substitution. However, this missense mutation was reported as a common polymorphism in the normal population in dbSNP. Also, this codon is not strictly conserved in other organisms with Ile being the most common amino acid at this position in mammalians. This indicates that p.Val120Ile is not causally related with FTLD.

Table 3: SNPs identified in *PCMT1.*

Region	Alias[1]	Genomic name[2]	cDNA name[3]	Protein name[4]	RefSNP ID[5]	Genotype Frequency[6]	Genotype frequency FTLD patients[7]
Intron 4	IVS4-79	g.43776C>T	c.298-79C>T		rs12182220	N.A.	0.348 CC 0.435 CT 0.217 TT
Exon 5	EX5+61	g.43915G>A	c.358G>A	p.Val120Ile	rs4816	0.390 GG 0.542 GA 0.068 AA	0.348 GG 0.435 GA 0.217 AA
Exon 8	**3'+85**	**g.61006C>T**	**c.*85C>T**			**N.A.**	**0.956 CC 0.044 CT**
Exon 8	3'+341	g.61262A>G	c.*341A>G		rs7818	0.400 AA 0.533 AG 0.067 GG	0.348 AA 0.435 AG 0.217 GG
Exon 8	3'+568	g.61489A>T	c.*568A>T		rs4552	0.292 AA 0.542 AT 0.167 TT	0.348 AA 0.435 AT 0.217 TT

Novel variant is shown in bold. [1] IVS: intron, EX: exon, 3': 3' UTR. [2] Numbering according to GenBank Accession Number NC_000006.10 and starting at nt 1. [3] Numbering according to the largest *PCMT1* transcipt (GenBank Accession Number NM_005389.1) and starting at translation initiation codon; * marks that the SNP is located in a non-coding part of the gene, the 3'UTR. [4] Numbering according to the largest *PCMT1* isoform (GenPept Accession Number NP_005380.1). [5] rs number according to the genome browser of the University of California Santa Cruz (UCSC). [6] European frequency as is known in the dbSNP database of NCBI; [7] Genotype frequency in the sequences of the 23 studied familial FTLD patients. When the sum is not one, we were unable to sequence all patients; N.A.: not applicable, frequencies were not known in the dbSNP of NCBI.

Since p.Val120Ile is not causally related to FTLD and we did not find a novel variant in a coding region, we can conclude that the cause of FTLD in these patients is not a mutation in the exons of *PCMT1*. However, we need to state that we were not able to amplify the first exon (exon 1) of this gene due to problems with the optimization of the PCR conditions. This

was most likely due to the high GC-content of this fragment (60.16%), a known issue for many first exons due to GC-rich regulatory elements. This problem might in some cases be solved by adding DMSO or Betaïne in the PCR reaction mixture. However, adding 10% DMSO did not improve amplification. This means that exon 1 might contain an unknown variant possibly causing FTLD. But this is only a hypothesis and still needs to be looked at.

3.2.2 Novel variants in *PXN*

The *PXN* gene is located on the long arm of chromosome 12 (12q24.31) and encodes a signal transduction adaptor protein. One of the major functions of *PXN* is the integration and dissemination of signals from growth factor receptors and integrins to effect efficient cellular migration [64]. PXN might be involved in GRN signal transduction. The eleven exons of *PXN* were sequenced. One SNP in exon 2 predicted a p.Ser73Gly substitution (Table 4). This missense mutation was a common polymorphism in the normal population in dbSNP. Serine is not strictly conserved in other organisms, glycine is also present at this position in other mammalians. This indicates that p.Ser73Gly is probably not causally related with FTLD.

Table 4: SNPs identified in the *PXN* gene.

Region	Alias[1]	Genomic name[2]	cDNA name[3]	Protein name[4]	RefSNP ID[5]	Genotype Frequency[6]	Genotype frequency FTLD patients[7]
Intron 1	IVS1-55	g.41276G>A	c.14-55G>A		rs11065054	0.933 GG 0.067 GA	0.956 GG 0.044GA
Exon 2	EX2+204	g.41533A>G	c.217A>G	p.Ser73Gly	rs4767884	0.708 GG 0.250 AG 0.042 AA	0.391 GG 0.304 AG 0.044 AA
Intron 2	**IVS2+58**	**g.41614G>T**	**c.240+58G>T**			**N.A.**	**0.957 GG 0.043 GT**
Exon 4	**EX4+23**	**g.42730G>A**	**c.379G>A**	**p.Ala127Ser**		**N.A.**	**0.957 GG 0.043 GA**
Exon 11	EX11+120	g.53246C>T	c.1527C>T	p.Tyr509	rs25664	0.958 CC 0.042 CT	0.957 CC 0.043 CT
Exon 11	3'+147	g.53540G>A	c.*1918G>A		rs3742039	0.150 GG 0.500 GA 0.350 AA	0.174 GG 0.565 GA 0.261 AA
Exon11	**3'+247**	**g.53640G>C**	**c.*2023G>C**			**N.A.**	**0.957 GG 0.043 GC**

Novel variants are shown in bold. [1] EX: exon, IVS: intron, 3': 3' UTR. [2] Numbering relative to the reverse complement of GenBank Accession Number NC_000012.10 and starting at nt 1. [3] Numbering according to the largest *PXN* transcipt (GenBank Accession Number NM_002859.1) and starting at translation initiation codon; * marks that the SNP is located in a non-coding part of the gene, the 3'UTR. [4] Numbering according to the largest *PXN* isoform (GenPept Accession Number NP_002850.1). [5] rs number according to the genome browser of the University of California Santa Cruz (UCSC). [6] European frequency as is known in the dbSNP database of NCBI; [7]

Genotype frequency in the sequences of the 23 studied familial FTLD patients. When the sum is not one, we were unable to sequence all the patients; N.A.: not applicable, frequencies were not known in the dbSNP of NCBI.

Because the first exon had a high GC-content (72.22%), we failed to optimize the PCR conditions for exon 1. So we are not able to draw conclusion about the whole gene. However, we identified three variants that were absent in the dbSNP database, p.Ala127Ser (Figure 4), g.53640G>C and g.41614G>T, which needed to be further investigated. p.Ala127Ser was located in the coding region of exon 4 and was of higher priority than the other two novel variants, which were located in the 3' UTR and in an intron. Transition of a GGC to an AGC on this location (p.Ala127Ser) resulted in a substitution of a highly conserved alanine to a serine.

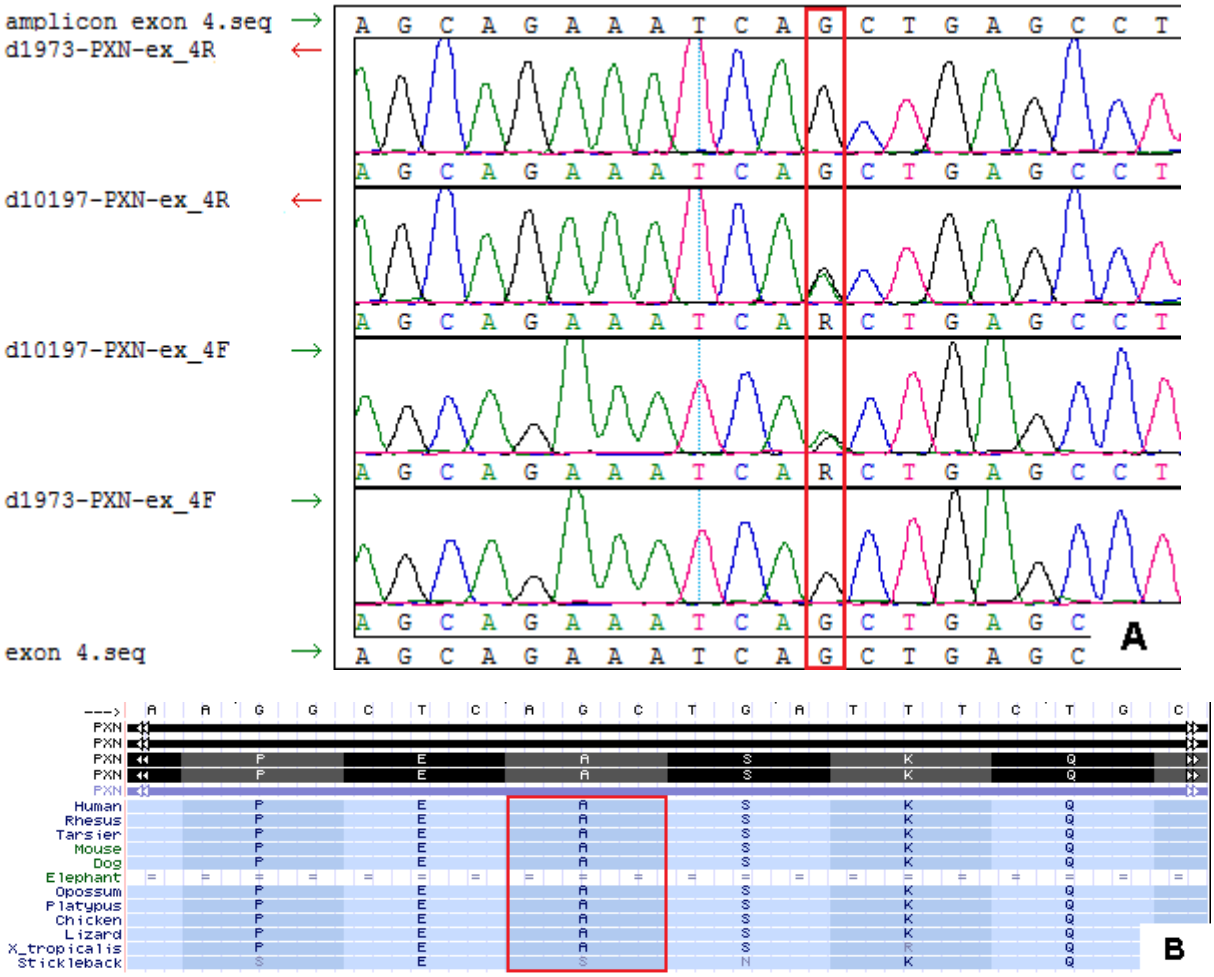


Figure 4: Novel variant in the *PXN* gene. A. Sequence of part of exon 4 of the *PXN* gene containing the p.Ala127Ser missense mutation. In patient sample d10197 a SNP can be seen which is located in exon 4 (EX4+23). Patient sample d1973 is used as an example of what it normally would be. **B.** Conservation plot with mainly alanine on position 127.

p.Ala127Ser could possibly be the cause of FTLD in patient d10197. This means that it is possible that this variation or SNP is inherited by the different family members and that they

develop FTLD because of the presence of this variation. However, to be sure that this is the case, we needed to check its absence in healthy individuals.

3.2.3 Presence of *PXN* p.Ala127Ser in control individuals

Since one of the 23 FTLD patients (d10197) showed a p.Ala127Ser mutation, *PXN* exon 4 was sequenced in 96 control individuals. Because the SNP was the clearest in the sequence of the reverse strand, only this one was sequenced. Analysis of the control individuals showed that two healthy individuals (d4803 and d5072) also had the transition from G to A on position EX4+23 in the *PXN* gene (Figure 5). As a consequence, it was concluded that this variation in exon 4 of *PXN* is not a causal variation in patient d10197. Thus, the presence of this variation does not lead to the development of a heritable form of FTLD.

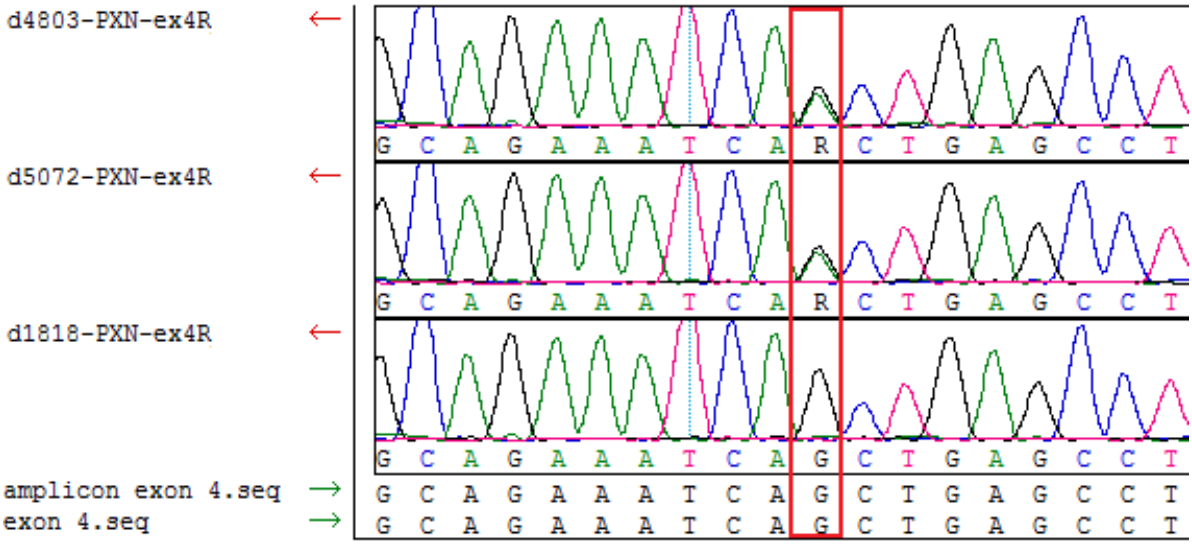


Figure 5: Sequence of exon 4 of *PXN* gene in control individuals. Both control d4803 as control d5702 have the SNP on position EX4+23. Sample d1818 shows the situation without the transition from a G to an A.

3.2.4 Novel variant in coding exon of *NUMA1*

NUMA1 is located on the long arm of chromosome 11 (11q13). NUMA1 is a protein that shuttles between the nuclear matrix in interphase and the spindle poles in mitosis. This protein appears to be important for the organization and stabilization of a focused array of microtubules at spindle poles [65, 66]. It might be correlated with the formation of spindle-shaped inclusions in FTLD-TDP. In this study, 25 of the 27 exons of *NUMA1* were sequenced. This was because we were not able to optimize the PCR conditions for exon 18 and 19 of *NUMA1.* Therefore we cannot make any conclusion for the complete *NUMA1* gene but only for the sequenced exons. Both the known and novel SNPs were noted together with information such as the rs number and the transition (Table 5).

Three missense mutations were identified in *NUMA1*, with one novel variant and two common polymorphisms in the normal population. The SNP in exon 21 predicted a p.Asp1795Asn amino acid substitution. Even though this missense mutation was reported as

a common polymorphism in the normal population in dbSNP, aspartic acid seems to be a highly conserved amino acid at this position in mammalians. So it could be that p.Asp1795Asn is causally related with FTLD.

Another missense mutation, located in exon 22, predicted a p.Tyr1836His amino acid substitution. This missense mutation was also reported to be a common polymorphism in the normal population in dbSNP. However, it seems to be that tyrosine is a highly conserved amino acid at this position in mammalians, which means that p.Tyr1836His could also be causally related with FTLD.

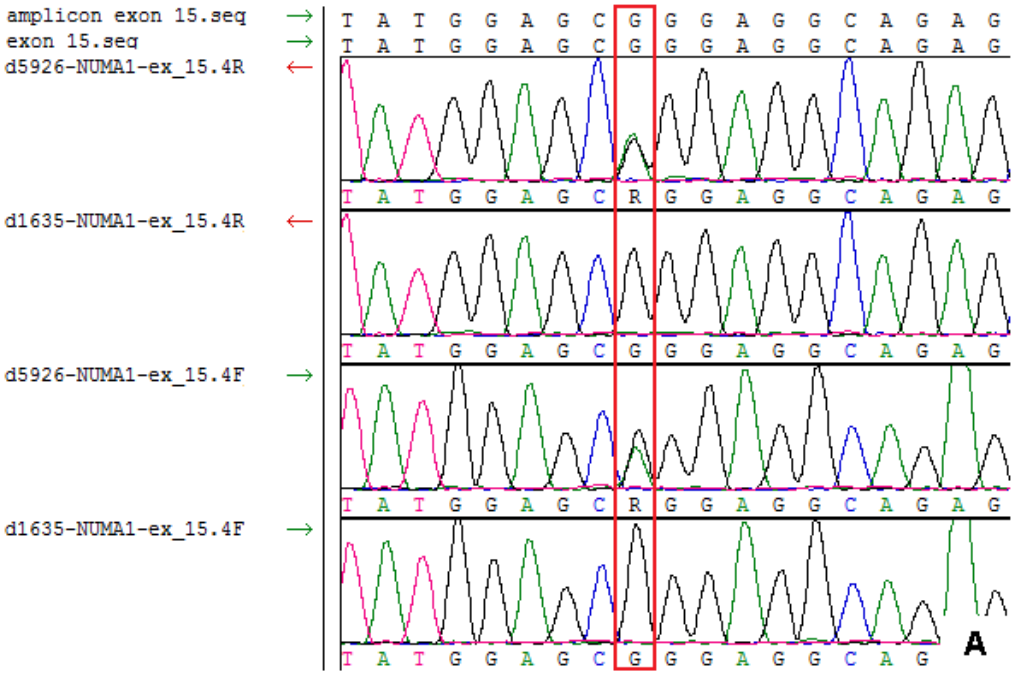
Table 5: SNPs in *NUMA1.*

Region	Alias[1]	Genomic name[2]	cDNA name[3]	Protein name[4]	RefSNP ID[5]	Genotype Frequency[6]	Genotype frequency FTLD patients[7]
Intron 2	IVS2+34	g.10720C>G	c.-33+34C>G		rs642573	0.750 CC 0.250 CG	0.043 CG 0.957 GG
Intron 3	IVS3-44>-43:insTG	g.51193_51194insTG	c.43-46_43-45insTG		rs10664166	1.000 insTG	1.000 insTG
Intron 5	IVS5-18	g.57363A>G	c.209-18A>G		rs2298457	0.333 AG 0.667 GG	0.043 AG 0.957 GG
Intron 6	IVS6-54	g.58055G>A	c.292-54G>A		rs10128658	0.138 GA 0.862 AA	0.043 GA 0.957 AA
Intron 8	IVS8-52	g.60859A>T	c.461-52A>T		rs2298456	0.750 AA 0.250 AT	0.043 AT 0.957 TT
Intron 10	IVS10-97	g.61918T>C	c.743-97T>C		rs949325	N.A.	0.043 TC 0.957 CC
Intron 12	IVS12-41>-37:delCGTTGA	g.62659_62664delCGTTGA	c.1151-41>-37delCGTTG		rs11278712	N.A.	0.043 WT DEL 0.957 DEL DEL
Intron 14	IVS14+32	g.64152T>G	c.1242+32T>G		rs1057992	0.167 TG 0.833 GG	0.043 TG 0.957 GG
Exon 15	**EX15+1673**	**g.65940G>A**	**c.2915G>A**	**p.Arg972 Gln**		**N.A.**	**0.957 GG 0.043 GA**
Intron 20	IVS20+115	g.71955C>A	c.5216+115C>A		rs2155145	0.017 CC 0.100 CA 0.883 AA	0.043 CA 0.956 AA
Exon 21	EX21+167	g.73259G>A	c.5383G>A	p.Asp1795 Asn	rs61744204	N.A.	0.913 CA 0.087 AA
Exon 21	EX21+184	g.73276G>A	c.5400G>A	p.Ser1800	rs61744213	N.A.	0.957 GG 0.043 GA

Exon 22	EX22+42	g.74307T>C	c.5506T>C	p.Tyr1836 His	rs355864 29	0.944 TT 0.056 TC	0.957 TT 0.043 TC
Intron 23	**IVS23-102**	**g.75610C>T**	**c.5830-102C>T**			**N.A.**	**0.957 CC** **0.043 CT**
Intron 25	IVS25+53	g.76356A>G	c.6123+53A>G		rs285236 5	0.200 AG 0.800 GG	0.043 AG 0.957 GG
Intron 25	IVS25+56	g.76359G>C	c.6123+56G>C		rs284585 7	0.700 GG 0.300 CC	0.043 GG 0.957 CC
Intron 26	IVS26-111	g.76879A>G	c.6337-111A>G		rs574368 2	0.950 AA 0.050 AG	0.957 AA 0.043 AG
Exon 27	3'+45	g.77046A>C	c.*45A>C		rs949323	0.130 AC 0.870 CC	0.043 AC 0.957 CC
Exon 27	3'+47	g.77048C>G	c.*47C>G		rs949324	0.300 CG 0.700 GG	0.043 CG 0.957 CG
Exon 27	3'+495	g.77496A>G	c.*495A>G		rs154130 4	0.117 AG 0.883 GG	0.043 AG 0.957 GG

Novel variants are shown in bold. [1] IVS: intron, EX: exon, 3': 3' UTR. [2] Numbering relative to the reverse complement of GenBank Accession Number NC_000011.8 and starting at nt 1. [3] Numbering according to the largest *NUMA1* transcipt (GenBank Accession Number NM_006185.2) and starting at translation initiation codon; * marks that the SNP is located in a non-coding part of the gene, the 3'UTR. [4] Numbering according to the largest *NUMA1* isoform (GenPept Accession Number NP_006176.2). [5] rs number according to the genome browser of the University of California Santa Cruz (UCSC). [6] frequency as is known in the dbSNP database of NCBI; [7] Genotype frequency in the sequences of the 23 studied familial FTLD patients. When the sum is not one, we were unable to sequence all the patients; N.A.: not applicable, frequencies were not known in the dbSNP of NCBI.

There were two novel variants found in *NUMA1* but only one was located in a coding region. p.Arg972Gln is located in exon 15 and resulted in a substitution of a highly conserved arginine to a glutamine at codon 972 (Figure 6).



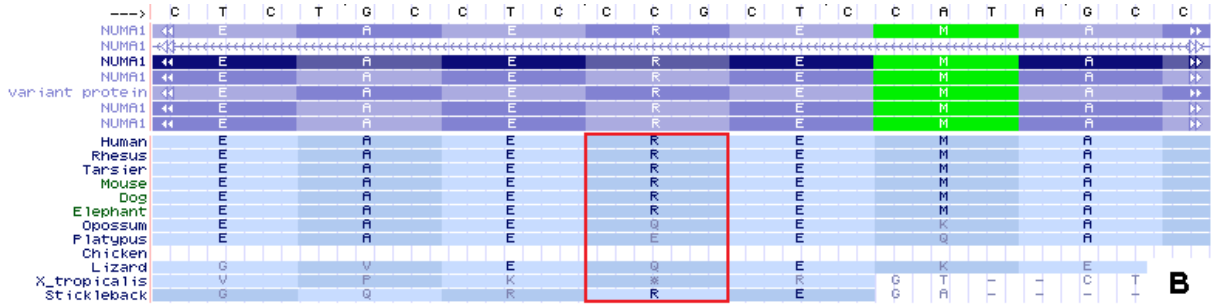


Figure 6: Sequence analysis *NUMA1*. A. Sequence of part of exon 15 of the *NUMA1* gene containing the p.Arg972Gln missense mutation. In patient sample d5926, a SNP can be seen which is located in exon 15 (EX15+1673). Patient sample d1635 is used as an example of what it normally would be. **B.** Conservation plot showing the highly conserved arginine at codon 972.

Since p.Arg972Gln is a transition from the highly conserved arginine in mammalians to a glutamine, p.Arg972Gln could be the cause of FTLD in patient d5926. However, this can only be stated if this SNP is absent in healthy individuals.

3.2.5 p.Arg972Gln in *NUMA1* exon 15 present in controls

Because one of the 23 FTLD patients (d5926) showed a p.Arg972Gln mutation, *NUMA1* exon 15 was sequenced in 96 control individuals. Since the SNP was the clearest in the sequence of the forward strand, only this one was sequenced. The analysis of the control individuals showed that healthy individuals (d3890, d1809 and d3860) also showed the p.Arg972Gln mutation in exon 15 of *NUMA1* (Figure 7). As a consequence, it could be concluded that p.Arg972Gln is not a causal variation in patient d5926. Thus, the presence of this variant does not mean that the carrier develops FTLD in a heritable form.

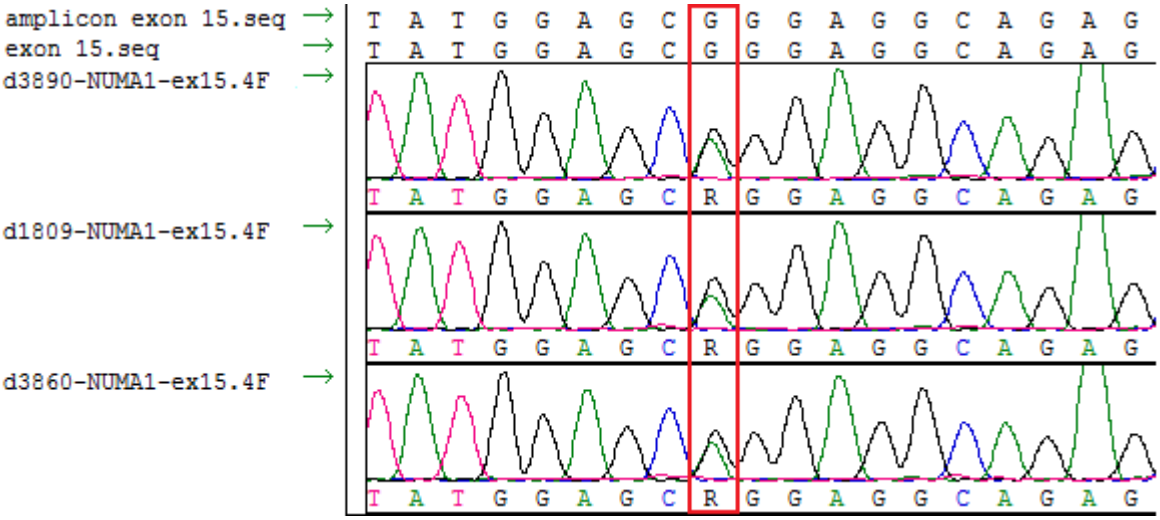


Figure 7: Sequence of part of exon 15 of the *NUMA1* gene in control individuals. Three controls (d3890, d1809 and d3860) also show the p.Arg972Gln mutation in exon 15.

3.2.6 Two intronic novel variants found in *CCNC*

All twelve coding exons of *CCNC* were screened for variations in the 23 FTLD patients. The *CCNC* gene is located on the long arm of chromosome 6 (6q21). *CCNC* encodes a protein that interacts with cyclin-dependent kinase 8 and induces the phosphorylation of the carboxy-terminal domain of the largest subunit of RNA polymerase II (MIM *123838). Since several proteins involved in RNA (e.g. *TARDBP* and *FUS*) processing are associated with MND, there might be a correlation between FTLD and CCNC. The variants with their characteristics are noted and can be seen in Table 6.

Table 6: SNPs in *CCNC*.

Region	Alias[1]	Genomic name[2]	cDNA name[3]	Protein name[4]	RefSNP ID[5]	Genotype Frequency[6]	Genotype frequency FTLD patients[7]
Intron 1	IVS1+32	g.351C>G	c.32+32C>G		rs41288985	N.A.	0.783 CC 0.217 CG
Intron 1	IVS1+132	g.451T>C	c.32+132T>C		rs330872	0.033 TT 0.400 TC 0.567 CC	0.043 TT 0.348 TC 0.609 CC
Intron 2	IVS2+64	g.6037delG	c.139+64delG		rs3831048	1.000 DEL G	0.100 DEL G
Intron 2	IVS2+115	g.6088C>G	c.139+115C>G		rs3736868	0.017 CC 0.400 CG 0.583 GG	0.783 CC 0.217 CG
Intron 2	**IVS2-73**	**g.7061G>A**	**c.140-73G>A**			**N.A.**	**0.957 GG 0.043 GA**
Intron 5	IVS5+129	g.10447A>C	c.346+129A>C		rs486881	0.034 AA 0.431 AC 0.534 CC	0.043 AA 0.348 AC 0.609 CC
Intron 6	IVS6-115	g.17672T>C	c.403-115T>C		rs2296155	N.A.	0.696 TT 0.301 TC
Intron 10	IVS10-82	g.23503G>T	c.679-82G>T		rs330806	0.317 GG 0.617 GT 0.067 TT	0.348 GG 0.435 GT 0.217 TT
Intron 11	**IVS11-93**	**g.25113C>T**	**c.798-93C>T**			**N.A.**	**0.870 CC 0.130 CT**
Exon 12	3'+575	g.25835A>G	c.*575A>G		rs1054227	0.600 AA 0.400 AG	0.652 AA 0.304 AG
Exon 12	3'+746	g.26006C>T	c.*746C>T		rs8366	0.667 CC 0.333 CT	0.652 CC 0.304 CT

Novel variants are shown in bold. [1] IVS: intron, EX: exon, 3': 3' UTR. [2] Numbering relative to the reverse complement of GenBank Accession Number NC_000006.10 and starting at nt 1. [3] Numbering according to the largest *CCNC* transcript (GenBank Accession Number NM_001013399.1) and starting at translation initiation codon; * marks that the SNP is located in a non-coding part of the gene, the 3'UTR. [4] Numbering according to the largest *CCNC* isoform (GenPept Accession Number NP_001013417.1). [5] rs number according to the genome

browser of the University of California Santa Cruz (UCSC). [6] frequency as is known in the dbSNP database of NCBI; [7] Genotype frequency in the sequences of the 23 studied familial FTLD patients. When the sum is not one, we were unable to sequence all the patients; N.A.: not applicable, frequencies were not known in the dbSNP of NCBI.

For the *CCNC* gene, two novel variants were found in the 23 FTLD patients from the Middelheim prospective study. However, both variants were located in an intron. As a result, we can exclude that a variation in this gene is responsible for FTLD in these patients.

3.2.7 Novel non-coding variants in *MED1*

MED1 is located on the long arm of chromosome 17 (17q12-q21.1) and encodes a protein which is, together with TRAP220, the main mediator subunit target for nuclear receptors. Examples of such receptors are the thyroid hormone, estrogen, vitamin D and retinoic acid receptors [67]. 16 of the 17 exons were sequenced because we were unable to optimize the PCR conditions for exon 8 of *MED1*. We identified six SNPs, including four novel variants (Table 7). However, none of the identified SNP were in a coding region of the gene. Three polymorphisms were located in introns and three in the 3' UTR of exon 17. So we can exclude that a variation in this gene is responsible for FTLD in these familial FTLD patients, unless exon 8 contains a novel coding polymorphism.

Table 7: SNPs identified in the *MED1* gene

Region	Alias[1]	Genomic name[2]	cDNA name[3]	Protein name[4]	RefSNP ID[5]	Frequency[6]	Genotype frequency FTLD patients[7]
Intron 0	**IVS0-69**		**c.-213-69T>C**			**N.A.**	**0.826 TT 0.130 TC**
Intron 0	IVS0-41		c.-213-41T>C		rs59110119	N.A.	0.826 CC 0.130 CG
Intron 12	**IVS12+10**		**c.852-10T>C**			**N.A.**	**0.869 TT 0.087 TC**
Exon 17	**3'+234**	**g.44034A>G**	**c.*234A>G**			**N.A.**	**0.956 AA 0.043 AG**
Exon 17	3'+2115	g.45915T>C	c.*2115T>C		rs6503513	0.458 TT 0.542 TC	0.609 TT 0.217 TC 0.043 CC
Exon 17	**3'+3169**	**g.46969A>G**	**c.*3169A>G**			**N.A.**	**0.869 AA 0.043 AG**

Novel variants are shown in bold. [1] IVS: intron, 3': 3' UTR. [2] Numbering relative to the reverse complement of GenBank Accession Number NC_000017.10 and starting at nt 1. [3] Numbering according to the largest *MED1* transcript (GenBank Accession Number NM_004774.3) and starting at translation initiation codon; * marks that the SNP is located in a non-coding part of the gene, the 3'UTR. [4] Numbering according to the largest *MED1* isoform (GenPept Accession Number NP_004765.2). [5] rs number according to the genome browser of the University of California Santa Cruz (UCSC). [6] frequency as is known in the dbSNP database of NCBI; [7] Genotype

frequency in the sequences of the 23 studied familial FTLD patients. When the sum is not one, we were unable to sequence all the patients; N.A.: not applicable, frequencies were not known in the dbSNP of NCBI.

3.2.8 15 novel variants in *TIAM1*

TIAM1 is located on the long arm of chromosome 21 (21q22.1). The protein TIAM1 is a member of the Db1 family of GNEFs that regulate small G proteins of the Rho family. Research has proven that TIAM1 plays an important role in cellular migration, invasion and/or adhesion. *TIAM1* is expressed in almost all adult tissues with a high expression in the brain and testis [68]. We identified 42 polymorphisms of which 15 were novel variants. 11 of the 15 novel polymorphisms are located in introns, three are located in the 3' UTR of the last exon and one SNP is located in the coding region of exon 18 (p.Asp1023Val; Table 8).

A SNP in exon 5 predicted a p.Thr109Ile amino acid substitution. Even though this missense mutation was reported as a common polymorphism in dbSNP, it seems that Thr is strictly conserved at this position in mammalians. This indicates that p.Thr109Ile could possibly be causally related with FTLD.

The other two variants in exon 5 predicted a p.Gly247Val. The resultant missense mutation was reported as a common polymorphism in the normal population in dbSNP. However, p.Gly247Val seems to be a substitution of a highly conserved glycine to a valine. Thus, it is possible that p.Gly247Val is also causally related with FTLD.

Table 8: SNPs identified in the *TIAM1* gene

Region	**Alias[1]**	**Genomic name[2]**	**cDNA name[3]**	**Protein name[4]**	**RefSNP ID[5]**	**Frequency [6]**	**Genotype frequency FTLD patients[7]**
Intron 1	IVS1-112	g.94831G>A	c.-421-112G>A		rs2252208	0.348 GG 0.391 GA 0.261 AA	0.304 GG 0.435 GA 0.261 AA
Exon 2	5'+30	g.94972C>T	c.-392C>T		rs9977815	0.379 CC 0.500 CT 0.121 TT	0.522 CC 0.304 CT 0.087 TT
Intron 2	**IVS2+28**	**g.95023C>A**	**c.-370+28C>A**			**N.A.**	**0.739 CC 0.043 CA**
Exon 5	EX5+326	g.292328C>T	c.326C>T	p.Thr109Ile	rs34882418	0.974 CC 0.026 CT	0.826 CC 0.087 CT
Exon 5	EX5+750	g.292741G>A	c.739G>A	p.Gly247Val	rs2070417	0.017 AA 0.203 AC 0.780 CC	 0.130 GA 0.826 GG
Exon 5	EX5+751	g.292742G>T	c.740G>T	p.Gly247Val	rs2070418	0.783 CC 0.200 CT 0.017 TT	0.826 GG 0.130 GT

Intron 6	**IVS6-33**	**g.313282A>G**	**c.1412-33A>G**			**N.A.**	**0.043 AG** **0.956 AA**
Intron 8	IVS8+39	g.333288T>A	c.1809+39T>A		rs2070416	0.792 AA 0.208 AT	0.869 TT 0.130 TA
Intron 9	IVS9+34	g.335603T>G	c.1995+34T>C		rs55980668	N.A.	0.913 TT 0.043 TG
Intron 9	**IVS9-78**	**g.341198G>A**	**c.1996-78G>A**			**N.A.**	**0.956 GG** **0.043 GA**
Intron 10	IVS10-21	g.345482G>A	c.2143-21G>A		rs2244187	N.A.	0.826 GG 0.174 GA
Intron 11	**IVS11+13**	**g.345590G>A**	**c.2217+13G>A**			**N.A.**	**0.956 GG** **0.043 GA**
Intron 12	IVS12-77	g.355886G>A	c.2389-77G>A		rs2833335	0.867 GG 0.133 GA	0.739 GG 0.174 GA
Intron 13	IVS13-180	g.363490T>C	c.2494-180C>T		rs2833331	0.208 TC 0.792 CC	0.130 CT 0.869 CC
Exon 14	EX14+59	g.363708G>C	c.2532G>C	p.Gln844His	rs16987932	1.000 CC	0.956 GG 0.043 GC
Intron 14	**IVS14+107**	**g.363858T>A**	**c.2575T>A**			**N.A.**	**0.956 TT** **0.043 TA**
Intron 14	**IVS14-218**	**g.371670T>A**	**c.2576-218T>A**			**N.A.**	**0.913 TT** **0.087 TA**
Intron 15	**IVS15+17**	**g.371991G>A**	**c.3134+17G>A**			**N.A.**	**0.043 GG** **0.956 GA**
Intron 17	IVS17-97	g.404450G>C	c.2992-97G>C		rs2833307	0.567 CC 0.417 CG 0.017 GG	 0.478 GC 0.521 GG
Intron 17	**IVS17-14**	**g.404533C>A**	**c.2992-14C>A**			**N.A.**	**0.912 CC** **0.087 CA**
Exon 18	**EX18+77**	**g.404623A>T**	**c.3068A>T**	**p.Asp1023Val**		**N.A.**	**0.956 AA** **0.043 AT**
Intron 18	IVS18+40	g.404766G>A	c.3171+40G>A		rs12482307	0.400 CC 0.467 CT 0.133 TT	0.522 GG 0.390 GA 0.043 AA
Intron 19	IVS19+39	g.405946C>T	c.3240+39C>T		rs723469	0.317 CC 0.500 CT 0.183 TT	0.348 CC 0.565 CT 0.087 TT
Intron 19	IVS19+82	g.405989T>G	c.3240+82T>G		rs723470	0.500 GG 0.433 GT 0.067 TT	0.696 GG 0.261 TG 0.043 TT
Intron 20	IVS20+102	g.406439T>C	c.3366+102T>G		rs9305453	N.A.	0.261 TC 0.739 CC
Intron 20	**IVS20-139**	**g.411831G>A**	**c.3367-139G>A**			**N.A.**	**0.956 GG** **0.043 GA**

Intron 21	**IVS21+78**	**g.412155G>T**	**c.3475+78G>T**			**N.A.**	**0.739 GG** **0.261 GT**
Intron 23	IVS23+113	g.417968T>C	c.3774+113T>G		rs2833301	0.833 AA 0.111 AG 0.056 GG	0.913 TT 0.087 TC
Intron 24	IVS24-76	g.427949T>C	c.3884-76T>C		rs2833289	0.762 AA 0.238 AG	0.913 TT 0.087 TC
Intron 26	IVS26-162	g.431659C>T	c.4046-162C>T		rs2070411	0.167 AA 0.542 AG 0.292 GG	0.348 TT 0.522 TC 0.130 CC
Intron 26	IVS26-107	g.431714T>C	c.4046-107T>G		rs2833286	0.792 AA 0.208 AG	0.913 TT 0.087 TC
Intron 26	IVS26-31	g.431790C>A	c.4046-31C>A		rs11700768	N.A.	0.913 CC 0.087 CA
Intron 26	IVS26-3	g.431818T>C	c.4046-3T>G		rs2070	0.087 AA 0.130 AG 0.783 GG	0.130 TT 0.522 TC 0.348 CC
Intron 28	**IVS28+20**	**g.434471G>A**	**c.4306+20G>A**			**N.A.**	**0.956 GG** **0.043 GA**
Intron 28	IVS28+130	g.434581T>C	c.4306+130T>C		rs2833282	0.600 AA 0.200 AG 0.200 GG	0.609 TT 0.348 TC 0.043 CC
Intron 28	IVS28-49	g.438087C>T	c.4307-49C>T		rs2073371	0.083 AA 0.500 AG 0.417 GG	0.348 TT 0.348 CT 0.304 CC
Exon 29	EX29+125	g.438260T>C	c.4431T>C	p.Gly1477	rs762194	0.345 AA 0.500 AG 0.155 GG	0.261 TT 0.391 TC 0.348 CC
Exon 29	**3'+16**	**g.438636C>T**	**c.*34C>T**			**N.A.**	**0.739 CC** **0.087 CT**
Exon 29	3'+194	g.438814A>G	c.*194A>G		rs2070409	0.167 CC 0.500 CT 0.333 TT	0.304 GG 0.261 AG 0.261 AA
Exon 29	3'+907	g.439526A>G	c.*907A>G		rs17551019	0.083 CT 0.917 TT	0.826 AA 0.130 AG
Exon 29	**3'+1603**	**g.440223G>A**	**c.*1603G>A**			**N.A.**	**0.956 GG** **0.043 GA**
Exon 29	**3'+1716**	**g.440336G>A**	**c.*1716G>A**			**N.A.**	**0.913 AA** **0.087 AG**

Novel variants are shown in bold. [1] IVS: intron, EX: exon, 3': 3' UTR. [2] Numbering relative to the reverse complement of GenBank Accession Number NC_000021.8 and starting at nt 1. [3] Numbering according to the largest *TIAM1* transcript (GenBank Accession Number NM_0035253.2) and starting at translation initiation codon; * marks that the SNP is located in a non-coding part of the gene, the 3'UTR. [4] Numbering according to the largest *TIAM1* isoform (GenPept Accession Number NP_003244.2). [5] rs number according to the genome browser of the University of California Santa Cruz (UCSC). [6] frequency as is known in the dbSNP database of NCBI; [7] Genotype

frequency in the sequences of the 23 studied familial FTLD patients. When the sum is not one, we were unable to sequence all the patients; N.A.: not applicable, frequencies were not known in the dbSNP of NCBI.

A SNP in exon 14 predicted a p.Gln844His amino acid substitution. This polymorphism was reported as a common polymorphism in the normal population in dbSNP. However, glutamine seems to be highly conserved at this position in other organisms which means that this amino acid substitution p.Gln844His could be causally related with FTLD.

There was also one novel variant present in the coding region of exon 18 (p.Asp1023Val, Figure 8). This amino acid substitution could be causally related to FTLD since aspartic acid seems to be highly conserved in other organisms, thus aspartic acid is the most common amino acid at this position in mammalians. However, to ensure that p.Asp1023Val causes FTLD, we needed to check the absence of this variant in healthy individuals. But, due to lack of time, we were unable to check for this polymorphism in the healthy individuals.

We were also unable to optimize the PCR conditions of the first exon of *TIAM1* due to the high GC content of this fragment (80.39%). Thus we cannot draw a final conclusion for the whole gene.

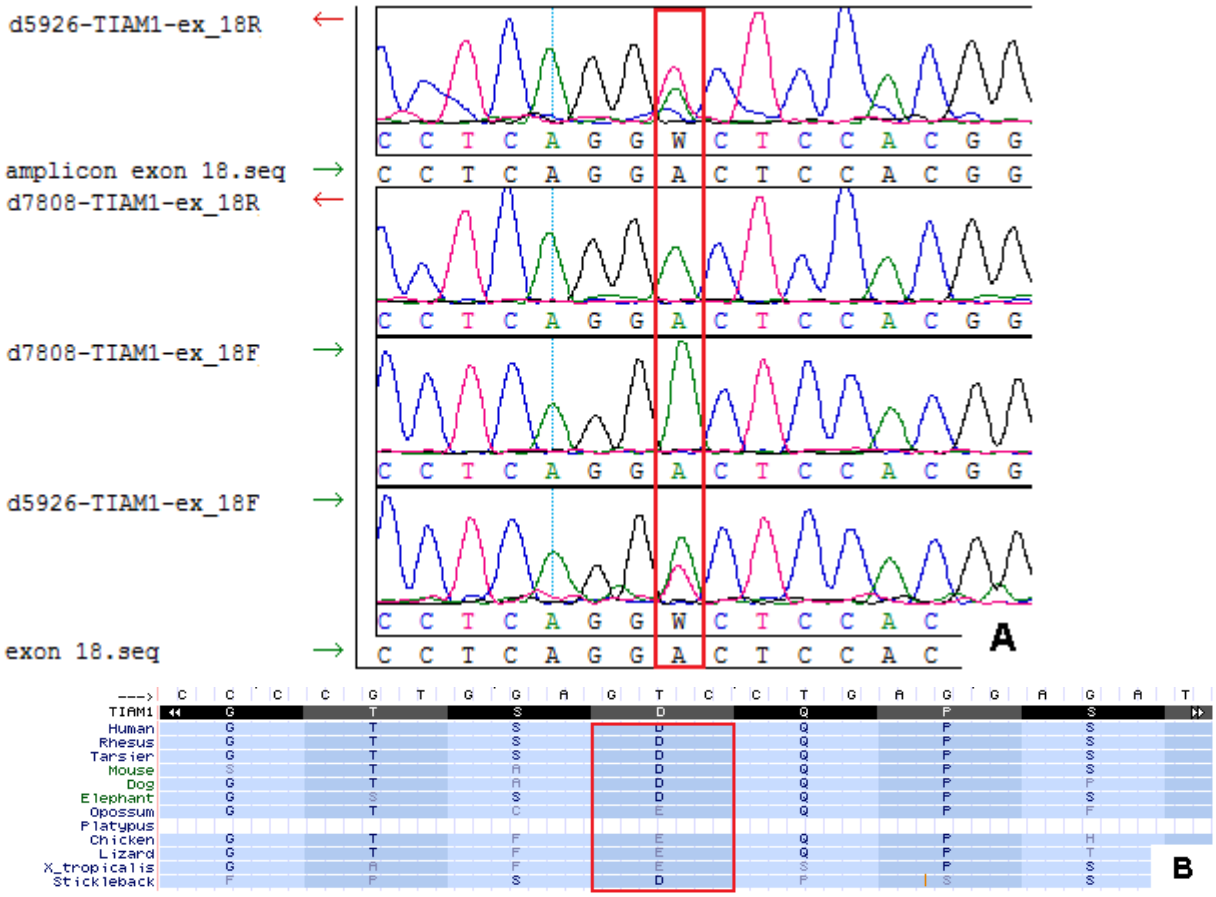


Figure 8: Sequence analysis of the *TIAM1* gene. A. Sequence of part of exon 18 of *TIAM1* containing the p.Asp1023Val missense mutation. In patient sample d5926, a SNP can be seen which is located in exon 18 (EX18+77). Patient sample d7808 is used as an example of what it normally would be. **B.** Conservation plot showing the conservation of aspartic acid in other mammalians.

3.2.9 Novel coding variant in *ATP2A2*

We sequenced 20 of the 21 exons of the *ATP2A2* gene. The first exon could not be sequenced because we were not able to optimize the PCR conditions. This was probably the result of the high GC content of this exon (71.89%). The *ATP2A2* gene is located on the long arm of chromosome 12 (12q23-q24.1) and encodes the sarco-endoplasmic reticulum calcium pumping ATPase type 2 (SERCA2) pump. This pump is extremely important in intracellular calcium signalling in the muscles [69]. We identified six polymorphisms of which one novel variant (Table 9).

Table 9: SNPs identified in the *ATP2A2* gene

Region	Alias[1]	Genomic name[2]	cDNA name[3]	Protein name[4]	RefSNP ID[5]	Frequency[6]	Genotype frequency FTLD patients[7]
Intron 3	IVS3-18	g.1537G>A	c.220-18G>A		rs35235621	N.A.	0.957 GG 0.043 AG
Intron 13	IVS13-81	g.63952A>G	c.1762-81A>G		rs11065631	0.817 AA 0.183 AG	0.869 AA 0.087 AG
Exon 15	EX15+75	g.65676G>A	c.2091G>A	p.Ala724	rs56243033	N.A.	0.652 GG 0.130 GA
Intron 15	IVS15-132	g.66474A>G	c.2319-132A>G		rs3026483	N.A.	0.869 AA 0.087 AG
Exon 17	**EX17+29**	**g.63688T>G**	**c.2550T>G**	**p.Ala850**		**N.A.**	**0.913 TT** **0.043 TG**
Intron 18	IVS18+54	g.68810G>A	c.2741+54G>A		rs1860561	0.034 AA 0.259 GA 0.707 GG	0.087 AA 0.435 GA 0.348 GG

Novel variant is shown in bold. [1] IVS: intron, EX: exon. [2] Numbering relative to GenBank Accession Number NC_000012.11 and starting at nt 1. [3] Numbering according to the largest *ATP2A2* transcript (GenBank Accession Number NM_170665.3) and starting at translation initiation codon; [4] Numbering according to the largest *ATP2A2* isoform (GenPept Accession Number NP_733765.1). [5] rs number according to the genome browser of the University of California Santa Cruz (UCSC). [6] frequency as is known in the dbSNP database of NCBI; [7] Genotype frequency in the sequences of the 23 studied familial FTLD patients. When the sum is not one, we were unable to sequence all the patients; N.A.: not applicable, frequencies were not known in the dbSNP of NCBI.

The mutation in exon 17 (p.Ala850) did not predict an amino acid substitution since it is located at the third “wobble” codon position. Thus p.Ala850 is called a silent mutation. This substitution does not result in a change at the protein level but can result in FTLD (Figure 9). To be sure that this silent mutation causes FTLD, we needed to check its absence in healthy individuals.

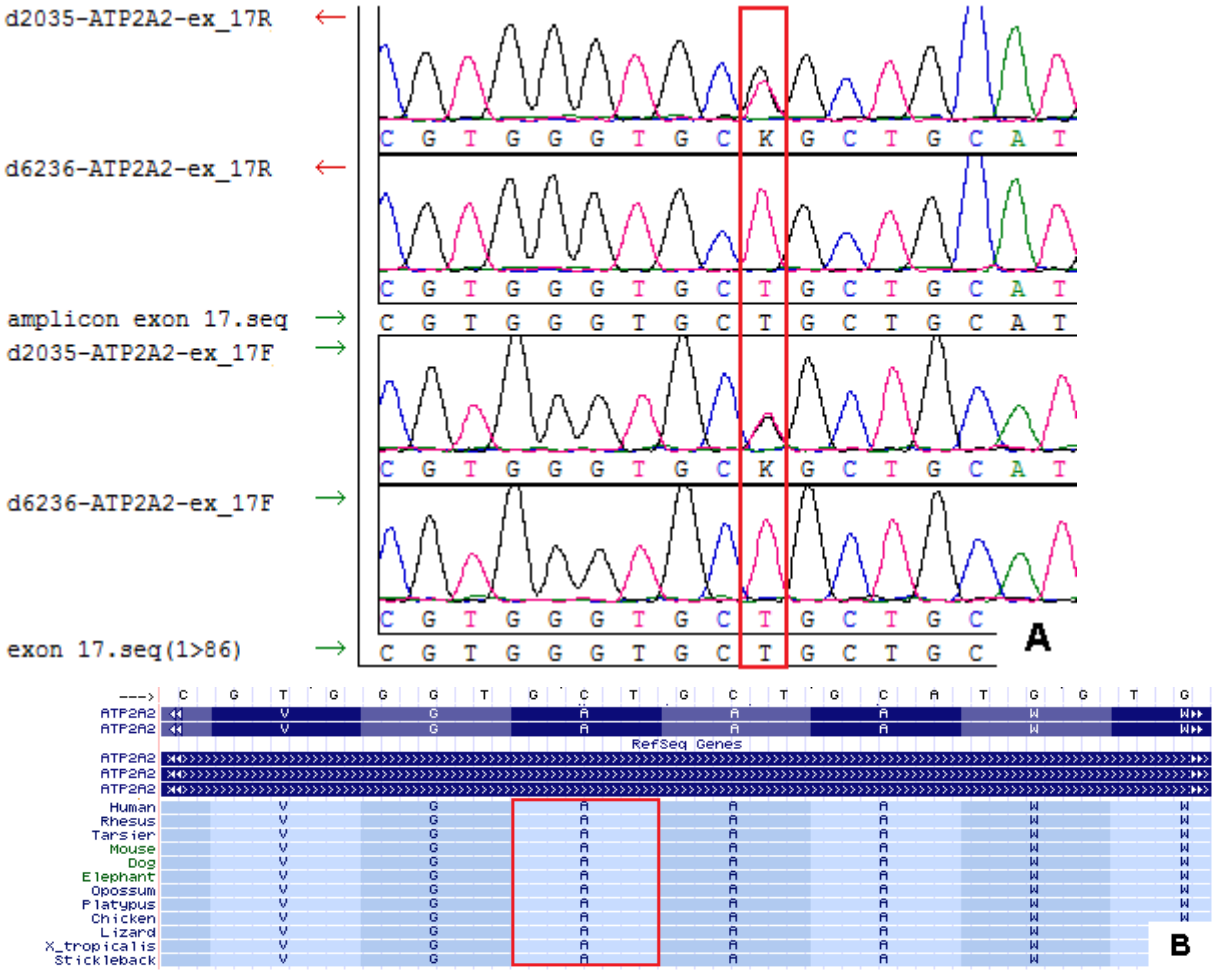


Figure 9: Analysis of *ATP2A2*. A. Sequence of part of exon 17 of *ATP2A2* containing the p.Ala850 silent mutation. In patient sample d2035, a SNP can be seen which is located in exon 17 (EX17+29). Patient sample d6236 is used as an example of what it normally would be. **B.** Conservation plot showing the highly conserved alanine on position 850.

3.2.10 Absence of p.Ala850 in 192 healthy individuals.

Since only one patient showed a p.Ala850 mutation, *ATP2A2* exon 17 was sequenced in 192 healthy individuals. However, none of the sequenced healthy individuals showed this variation. It is thus possible that this variant is causally related with FTLD. Silent mutations may have a lot of functional consequences and may lead to a disease state by disrupting exonic splicing enhancers or by changing the substrate specificity [70, 71]. Which effect p.Ala850 has in FTLD still needs to be elucidated.

4 SYNTHESIS

FTLD is the second most common cause of dementia in people under the age of 65 years. About 30-50% of FTLD patients have a positive family history of dementia which suggests that genetics play a major role in the etiology of FTLD. For almost 50% of these inherited forms of FTLD, mutations are found in *GRN*, *VCP*, *CHMP2B* and *MAPT*. For the other 50% of familial FTLD patients, no causal mutation is found. As a result, we hypothesized that other causal genes and risk variants that are active in a common pathway shared with the known FTLD-TDP genes *GRN*, *VCP* and *TARDBP* could be found.

We selected eight different genes (*PCMT1*, *PXN*, *NUMA1*, *CCNC*, *MED1*, *TIAM1, BCL2* and *ATP2A2*) that all showed a direct or indirect interaction with one of the FTLD-TDP genes. Next we sequenced these eight genes on the genomic DNA of 23 FTLD patients derived from the Middelheim prospective study of dementia.

In *PCMT1*, we identified five polymorphisms, four reported variants and one novel SNP. The novel variant was located in the 3' UTR of the last exon of *PCMT1* and we could not check the absence of this variant in healthy individuals due to a lack of time. The *PCMT1* gene also contained a reported missense mutation (p.Val120Ile). Since Ile was the most common amino acid at this position in mammalians, we concluded that p.Val120Ile is not causally related with FTLD. However, it could be that this missense mutation is associated with a higher risk of developing FTLD. No data are available about the role of *PCMT1* in neurodegenerative diseases such as FTLD. However, there are research groups which looked at the *PCMT1* gene and the reported missense mutation in other diseases such as spina bifida. They found that the Val/Val genotype was associated with a reduction in risk for spina bifida [72]. However, to conclusively investigate the correlation between *PCMT1* and FTLD, we need to perform association studies and look for the novel variation (g.61006C>T) in a large set of FTLD patients and neurologically healthy control individuals.

The next gene we looked at was *PXN.* We identified five reported polymorphisms and two novel ones. Again, one novel variant (g.53640G>C) was located in the 3' UTR of the last exon and the absence of this SNP still needs to be checked in healthy individuals. The other novel variant was a missense mutation (p.Ala127Ser) and we checked for the absence in neurologically healthy individuals. This polymorphism was not strictly associated with FTLD patients since two healthy individuals carried the missense mutation. However, the possibility exists again that this missense mutation is associated with a higher risk on developing FTLD. To be sure whether this statement is true, an association study needs to be carried out. The results of this association study is very important since *PXN* is already reported in another neurodegenerative disorder, namely AD. Grace et al [73] found that fibrillar Aβ binds to

integrin receptors and induces their clustering. This clustering activates the paxillin protein which results in a cascade of actions. One of the results of these actions is the hyperphosphorylation of tau and thus the destabilization of the microtubules and neuronal dystrophy. Thus, according to Grace et al [73], *PXN* seems to play an important role in neurodegeneration.

A third gene, *NUMA1,* encodes a protein which can only be found in the nucleus during interphase in mitotic cells. However, neurons are terminally post-mitotic cells and display the NUMA1 protein both in the nucleus and in the cytoplasm [74]. We identified two novel variants of which one was located in the coding region of exon 15 (p.Arg972Gln) and the other in intron 23 (g.75610C>T). We controlled for the presence of p.Arg972Gln in neurologically healthy control individuals but the variation was present in three healthy individuals. Nevertheless, this does not mean that p.Arg972Gln is not associated with a higher risk on developing FTLD. This still needs to be verified because Ferhat et al [74] provide information about a potentially important role for NUMA1 in the organization of microtubules in the neurons. When there is a problem with this microtubule organization, neuronal dystrophy occurs.

After sequencing the *CCNC* gene, no variant was found in a coding region of the gene. However, we found nine intronic polymorphisms of which two were novel variants and two reported SNPs in the 3' UTR region of exon 12. These novel intronic variations can be important because they could influence the splicing process which could possibly generate a different protein than the splicing process normally would. To my knowledge, the *CCNC* gene is not associated with FTLD in literature. However, Jiang et al [75] proved that there was an upregulation of *CCNC* in subpopulations of residual motor neurons in ALS. They also found that the upregulation of *CCNC* is a relatively late event which leads to neuronal death. This finding can be very important because FTLD can also be associated with ALS or motor neuron disease. So it is possible that *CCNC* plays a role in FTLD-ALS. However, in this case, it would be better to screen FTLD-ALS patients for mutations in *CCNC*.

In the *MED1* gene, we identified four novel non-coding variants. Two were located in introns and the other two were located in the 3' UTR. All four variants can have an important role in the splicing process. MED1 is a protein belonging to the Mediator complex which plays a key coregulatory role in eukaryotic transcription by RNA polymerase II. As for the other discussed genes so far, no literature is available that provides information about the role of *MED1* in FTLD. However, it has been shown that a MED1 binding site is present in the promoter of *APOE*. *APOE* is important in AD and researchers also have looked at its role in FTLD. So a mutation in *MED1* could lead to problems with the activation of transcription of the genes which depend on the activity of MED1.

One of the genes we selected for analysis but which was not present in our 'Results and Discussion' section was *BCL2* (18q21.33). During the period of this internship, we were able to find the optimal PCR conditions for all the exons of *BCL2* but the sequencing of all the different parts of the third exon gave difficulties. As a result, we cannot say that the *BCL2* gene is not important in the etiology of FTLD. *BCL2* belongs to a family of genes that are responsible for controlling cell homeostatic processes. Some of these genes promote programmed cell death or apoptosis and others seem to have a negative influence on apoptosis [76]. In neurodegenerative disorders, it is believed that apoptosis plays an important role. As a result, *BCL2* could play a major role in neurodegeneration.

TIAM1 is a gene that regulates small G proteins and is of critical importance to cellular function. *TIAM1* is highly expressed in brain and testis and may play a role in embryogenesis [68]. Analysis of the *TIAM1* gene resulted in a list of 42 polymorphisms with 15 novel variations. Here, we identified three novel variations located in the 3' UTR of the last exon of *TIAM1*, eleven novel variants located in introns spread over the whole gene and one novel missense mutation in exon 18 (p.Asp1023Val). Presence of this missense mutation still needs to be screened in healthy individuals to be able to conclude whether it is important in the etiology of FTLD or not. We were not able to perform this analysis yet. So the possibility exist that this missense mutation is the cause of FTLD. *TIAM1* is not yet linked to FTLD but has been looked at in AD. Mendoza-Naranjo et al [77] showed a upregulation of *TIAM1* after stimulation with Aβ in AD. This upregulation of *TIAM1* would result in increased activation of RAC1. The end result of this cascade is, according to Mendoza-Naranjo et al, actin cytoskeletal alterations as are seen in AD. Thus it seems that *TIAM1* plays a role in neurodegeneration but its role in FTLD still needs to be proven.

The last gene we analyzed is *ATP2A2*. After mutational sequencing, we identified six polymorphisms with one novel variant (p.Ala850). We looked for the presence of this variant in 192 healthy individuals but none of them carried this polymorphism. Now, it does not seem logical that a gene which encodes a SERCA2 pump (a Ca^{2+} pump of the sarcoplasmatic reticulum in muscle) is analyzed in a neurological disease as FTLD. Though, earlier reports have linked mutations in this gene to bipolar disorder, another neurological disorder. However, Jacobsen et al [78] recently proved that there was no evidence for the involvement of *ATP2A2* in producing susceptibility to bipolar disorder. So the link between *ATP2A2* and neurological disorders is still unclear. p.Ala850 is a silent mutation and as a result, does not change the protein. Nevertheless, this silent mutation may have different effects on the mRNA processing. For instance, in phenylketonuria, a silent mutation is reported inducing exon skipping in the phenylalanine hydroxylase gene [70]. This exon skipping resulted in the loss of exon 11 on mRNA level and consequently in phenylketonuria. Kimchi-Sarfaty et al [71] reported a silent mutation in the *MDR1* gene resulting in alteration of the substrate

specificity. Consequently, p.Ala850 may have an effect on the mRNA of the *ATP2A2* gene and result in FTLD.

For all the eight selected genes it is important that all the analyses are completed. This means that all the exons for which we failed to find the optimized PCR conditions, need to be sequenced. These exons are the first exons of *PCMT1*, *PXN*, *TIAM1* and *ATP2A2*. Exon 8 of *MED1* and exon 18 and 19 of *NUMA1* also still need to be optimized and sequenced. As mentioned before, the whole *BCL2* gene must be sequenced and analyzed. It is also important that we still check the absence of some novel variations in healthy individuals (Table 10). These novel variants also include variants located in the 3' UTR of a gene. Mutations in this region may have important effects since they can influence the binding affinity of miRNAs. When there is a higher affinity for miRNA, mRNA is degraded and translational arrest can occur. As a result, it is very important that these novel variants located in the 3' UTR of a gene are sequenced on the genomic DNA of healthy individuals.

Table 10: Novel variations that still need to be checked in healthy individuals.

Gene	Region	Name
PCMT1	Exon 8	g.61006C>T
PXN	Exon 11	g.53640G>C
MED1	Exon 17	g.44034A>G
MED1	Exon 17	g.46969A>G
TIAM1	Exon 18	p.Asp1023Val
TIAM1	Exon 29	g.438636C>T
TIAM1	Exon 29	g.440223C>A
TIAM1	Exon 29	g.440336G>A
ATP2A2	Exon 17	p.Ala850

When one of these novel polymorphisms is absent in a sufficient number of healthy individuals, functional analysis needs to be performed. This means that we need to generate a transgenic neural cell line with the mutated variant, for example p.Asp1023Val in the *TIAM1* gene, and look what the effect of this variant would be on the processing of *TARDBP*. When we have sufficient data on the effect of the novel variant on *TARDBP* processing in vitro, we can check if our findings are the same in a mouse model. But this is only when we are very certain that the novel variant is important in the etiology of FTLD.

In conclusion, this research resulted in several possible causal FTLD pathways (Figure10). One possible pathway involves *ATP2A2*. This pathway can work in different ways, for instance *GRN* can activate *ATP2A2* directly or can first activate *VCP* through *CFTR* or *Akt1* and then *VCP* can activate *ATP2A2*. This *ATP2A2* can activate other genes which result in the activation of *TARDBP*. Another possible pathway involves activation of *TIAM1*. *GRN* directly activates *TIAM1* and *TIAM1* activates other genes such *DICER* which results in the

activation of *TARDBP*. This is another possible pathway of FTLD. However, since we were not able to sequence *BCL2*, we cannot say that it is not involved in the pathway which causes FTLD. *BCL2* can be directly activated by *GRN* or indirectly by *VCP* (*VCP – CFTR – TIAM1 – MYC – BCL2*). *BCL2* in turn activates *CCNC* and then activates *TARDBP*.
Which of these pathways is involved in the etiology of FTLD or if another pathway must exist still needs to be elucidated. Future research will find an answer to this question.

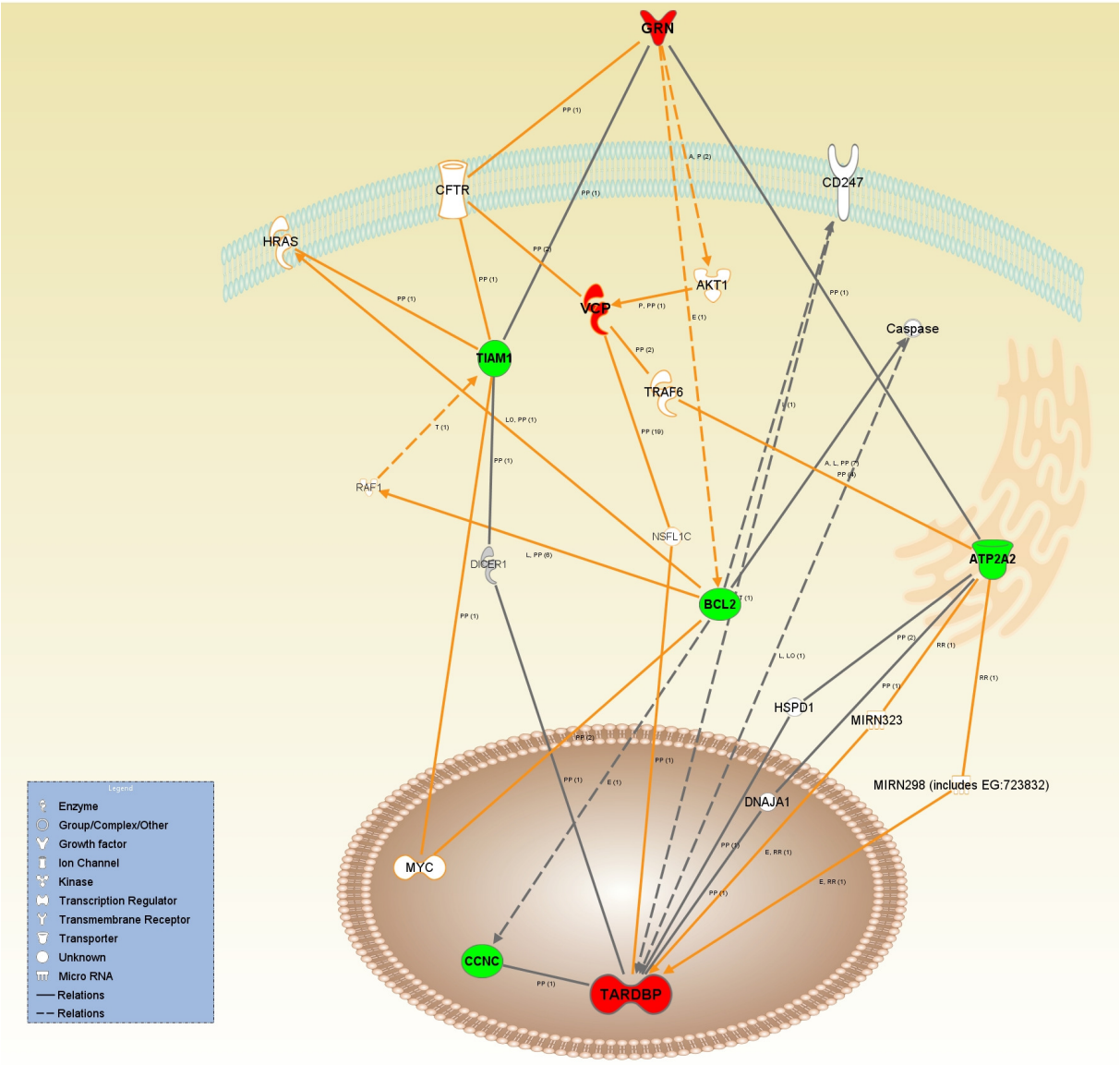


Figure 10: Conclusion in which the three possible pathways of FTLD are shown. The FTLD-TDP genes are shown in red and the sequenced genes in green.

5 REFERENCES

1. Ferri, C.P., et al., *Global prevalence of dementia: a Delphi consensus study.* Lancet, 2005. **366**(9503): p. 2112-7.
2. Association, A.s., *2009 Alzheimer's disease Facts and Figures.* Alzheimer's & Dementia, 2009. **5**(3): p. 1-80.
3. Ratnavalli, E., et al., *The prevalence of frontotemporal dementia.* Neurology, 2002. **58**(11): p. 1615-21.
4. Harvey, R.J., M. Skelton-Robinson, and M.N. Rossor, *The prevalence and causes of dementia in people under the age of 65 years.* J Neurol Neurosurg Psychiatry, 2003. **74**(9): p. 1206-9.
5. Kumar-Singh, S. and C. Van Broeckhoven, *Frontotemporal lobar degeneration: current concepts in the light of recent advances.* Brain Pathol, 2007. **17**(1): p. 104-14.
6. Neary, D., J.S. Snowden, and D.M. Mann, *Classification and description of frontotemporal dementias.* Ann N Y Acad Sci, 2000. **920**: p. 46-51.
7. Sleegers, K., et al., *Molecular pathogenesis of frontotemporal lobar degeneration: basic science seminar in neurology.* Arch Neurol, 2008. **65**(6): p. 700-4.
8. Neary, D., et al., *Frontotemporal lobar degeneration: a consensus on clinical diagnostic criteria.* Neurology, 1998. **51**(6): p. 1546-54.
9. Sampathu, D.M., et al., *Pathological heterogeneity of frontotemporal lobar degeneration with ubiquitin-positive inclusions delineated by ubiquitin immunohistochemistry and novel monoclonal antibodies.* Am J Pathol, 2006. **169**(4): p. 1343-52.
10. Neumann, M., et al., *Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis.* Science, 2006. **314**(5796): p. 130-3.
11. Josephs, K.A., et al., *Frontotemporal lobar degeneration with ubiquitin-positive, but TDP-43-negative inclusions.* Acta Neuropathol, 2008. **116**(2): p. 159-67.
12. Mackenzie, I.R., et al., *Atypical frontotemporal lobar degeneration with ubiquitin-positive, TDP-43-negative neuronal inclusions.* Brain, 2008. **131**(Pt 5): p. 1282-93.
13. Mackenzie, I.R., et al., *Nomenclature for neuropathologic subtypes of frontotemporal lobar degeneration: consensus recommendations.* Acta Neuropathol, 2009. **117**(1): p. 15-8.
14. Mackenzie, I.R. and R. Rademakers, *The molecular genetics and neuropathology of frontotemporal lobar degeneration: recent developments.* Neurogenetics, 2007. **8**(4): p. 237-48.

15. Dickson, D.W., et al., *Office of Rare Diseases neuropathologic criteria for corticobasal degeneration.* J Neuropathol Exp Neurol, 2002. **61**(11): p. 935-46.
16. Hauw, J.J., et al., *Preliminary NINDS neuropathologic criteria for Steele-Richardson-Olszewski syndrome (progressive supranuclear palsy).* Neurology, 1994. **44**(11): p. 2015-9.
17. Rowland, L.P. and N.A. Shneider, *Amyotrophic lateral sclerosis.* N Engl J Med, 2001. **344**(22): p. 1688-700.
18. Van Deerlin, V.M., et al., *Familial frontotemporal dementia: from gene discovery to clinical molecular diagnostics.* Clin Chem, 2003. **49**(10): p. 1717-25.
19. van Swieten, J.C. and P. Heutink, *Mutations in progranulin (GRN) within the spectrum of clinical and pathological phenotypes of frontotemporal dementia.* Lancet Neurol, 2008. **7**(10): p. 965-74.
20. van der Zee, J., et al., *Frontotemporal lobar degeneration with ubiquitin-positive inclusions: a molecular genetic update.* Neurodegener Dis, 2007. **4**(2-3): p. 227-35.
21. Hutton, M., et al., *Association of missense and 5'-splice-site mutations in tau with the inherited dementia FTDP-17.* Nature, 1998. **393**(6686): p. 702-5.
22. Binder, L.I., A. Frankfurter, and L.I. Rebhun, *The distribution of tau in the mammalian central nervous system.* J Cell Biol, 1985. **101**(4): p. 1371-8.
23. Gu, Y., F. Oyama, and Y. Ihara, *Tau is widely expressed in rat tissues.* J Neurochem, 1996. **67**(3): p. 1235-44.
24. Bugiani, O., *The many ways to frontotemporal degeneration and beyond.* Neurol Sci, 2007. **28**(5): p. 241-4.
25. Rademakers, R., M. Cruts, and C. van Broeckhoven, *The role of tau (MAPT) in frontotemporal dementia and related tauopathies.* Hum Mutat, 2004. **24**(4): p. 277-95.
26. Mutations, *Alzheimer Disease & Frontotemporal Dementia Mutation Database.* http://www.molgen.ua.ac.be/FTDMutations, 5 February 2009.
27. Trinczek, B., et al., *Tau regulates the attachment/detachment but not the speed of motors in microtubule-dependent transport of single vesicles and organelles.* J Cell Sci, 1999. **112 (Pt 14)**: p. 2355-67.
28. He, Z. and A. Bateman, *Progranulin (granulin-epithelin precursor, PC-cell-derived growth factor, acrogranin) mediates tissue repair and tumorigenesis.* J Mol Med, 2003. **81**(10): p. 600-12.
29. He, Z., et al., *Progranulin is a mediator of the wound response.* Nat Med, 2003. **9**(2): p. 225-9.
30. Cruts, M., et al., *Null mutations in progranulin cause ubiquitin-positive frontotemporal dementia linked to chromosome 17q21.* Nature, 2006. **442**(7105): p. 920-4.

31. Baker, M., et al., *Mutations in progranulin cause tau-negative frontotemporal dementia linked to chromosome 17.* Nature, 2006. **442**(7105): p. 916-9.
32. Maquat, L.E., *Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics.* Nat Rev Mol Cell Biol, 2004. **5**(2): p. 89-99.
33. Wang, Q., C. Song, and C.C. Li, *Molecular perspectives on p97-VCP: progress in understanding its structure and diverse biological functions.* J Struct Biol, 2004. **146**(1-2): p. 44-57.
34. Woodman, P.G., *p97, a protein coping with multiple identities.* J Cell Sci, 2003. **116**(Pt 21): p. 4283-90.
35. Dai, R.M. and C.C. Li, *Valosin-containing protein is a multi-ubiquitin chain-targeting factor required in ubiquitin-proteasome degradation.* Nat Cell Biol, 2001. **3**(8): p. 740-4.
36. Rabinovich, E., et al., *AAA-ATPase p97/Cdc48p, a cytosolic chaperone required for endoplasmic reticulum-associated protein degradation.* Mol Cell Biol, 2002. **22**(2): p. 626-34.
37. Guinto, J.B., et al., *Valosin-containing protein and the pathogenesis of frontotemporal dementia associated with inclusion body myopathy.* Acta Neuropathol, 2007. **114**(1): p. 55-61.
38. Kovach, M.J., et al., *Clinical delineation and localization to chromosome 9p13.3-p12 of a unique dominant disorder in four families: hereditary inclusion body myopathy, Paget disease of bone, and frontotemporal dementia.* Mol Genet Metab, 2001. **74**(4): p. 458-75.
39. Kimonis, V.E. and G.D. Watts, *Autosomal dominant inclusion body myopathy, Paget disease of bone, and frontotemporal dementia.* Alzheimer Dis Assoc Disord, 2005. **19 Suppl 1**: p. S44-7.
40. Brown, J., et al., *Familial non-specific dementia maps to chromosome 3.* Hum Mol Genet, 1995. **4**(9): p. 1625-8.
41. Cannon, A., et al., *CHMP2B mutations are not a common cause of frontotemporal lobar degeneration.* Neurosci Lett, 2006. **398**(1-2): p. 83-4.
42. Rizzu, P., et al., *CHMP2B mutations are not a cause of dementia in Dutch patients with familial and sporadic frontotemporal dementia.* Am J Med Genet B Neuropsychiatr Genet, 2006. **141B**(8): p. 944-6.
43. Sun, Y., et al., *VEGF-induced neuroprotection, neurogenesis, and angiogenesis after focal cerebral ischemia.* J Clin Invest, 2003. **111**(12): p. 1843-51.
44. Kalaria, R.N., et al., *Vascular endothelial growth factor in Alzheimer's disease and experimental cerebral ischemia.* Brain Res Mol Brain Res, 1998. **62**(1): p. 101-5.

45. Borroni, B., et al., *Preliminary evidence that VEGF genetic variability confers susceptibility to frontotemporal lobar degeneration.* Rejuvenation Res, 2008. **11**(4): p. 773-80.
46. Saunders, A.M., et al., *Association of apolipoprotein E allele epsilon 4 with late-onset familial and sporadic Alzheimer's disease.* Neurology, 1993. **43**(8): p. 1467-72.
47. Srinivasan, R., et al., *The apolipoprotein E epsilon4 allele selectively increases the risk of frontotemporal lobar degeneration in males.* J Neurol Neurosurg Psychiatry, 2006. **77**(2): p. 154-8.
48. Rosso, S.M., et al., *Apolipoprotein E4 in the temporal variant of frontotemporal dementia.* J Neurol Neurosurg Psychiatry, 2002. **72**(6): p. 820.
49. Short, R.A., et al., *Differences in tau and apolipoprotein E polymorphism frequencies in sporadic frontotemporal lobar degeneration syndromes.* Arch Neurol, 2002. **59**(4): p. 611-5.
50. Pickering-Brown, S.M., et al., *Apolipoprotein E epsilon4 allele has no effect on age at onset or duration of disease in cases of frontotemporal dementia with pick- or microvacuolar-type histology.* Exp Neurol, 2000. **163**(2): p. 452-6.
51. Thorns, V., L. Hansen, and E. Masliah, *nNOS expressing neurons in the entorhinal cortex and hippocampus are affected in patients with Alzheimer's disease.* Exp Neurol, 1998. **150**(1): p. 14-20.
52. Venturelli, E., et al., *Neuronal nitric oxide synthase C276T polymorphism increases the risk for frontotemporal lobar degeneration.* Eur J Neurol, 2008. **15**(1): p. 77-81.
53. Petersen, R.B., et al., *Signal transduction cascades associated with oxidative stress in Alzheimer's disease.* J Alzheimers Dis, 2007. **11**(2): p. 143-52.
54. Baker, M., et al., *Association of an extended haplotype in the tau gene with progressive supranuclear palsy.* Hum Mol Genet, 1999. **8**(4): p. 711-5.
55. Conrad, C., et al., *Genetic evidence for the involvement of tau in progressive supranuclear palsy.* Ann Neurol, 1997. **41**(2): p. 277-81.
56. Holm, I.E., et al., *A reassessment of the neuropathology of frontotemporal dementia linked to chromosome 3.* J Neuropathol Exp Neurol, 2007. **66**(10): p. 884-91.
57. Ayala, Y.M., et al., *Human, Drosophila, and C.elegans TDP43: nucleic acid binding properties and splicing regulatory function.* J Mol Biol, 2005. **348**(3): p. 575-88.
58. Neumann, M., et al., *TDP-43-positive white matter pathology in frontotemporal lobar degeneration with ubiquitin-positive inclusions.* J Neuropathol Exp Neurol, 2007. **66**(3): p. 177-83.
59. Schwab, C., et al., *Colocalization of transactivation-responsive DNA-binding protein 43 and huntingtin in inclusions of Huntington disease.* J Neuropathol Exp Neurol, 2008. **67**(12): p. 1159-65.

60. Kwong, L.K., et al., *TDP-43 proteinopathies: neurodegenerative protein misfolding diseases without amyloidosis.* Neurosignals, 2008. **16**(1): p. 41-51.
61. Engelborghs, S., et al., *Prospective Belgian study of neurodegenerative and vascular dementia: APOE genotype effects.* J Neurol Neurosurg Psychiatry, 2003. **74**(8): p. 1148-51.
62. Clarke, S., *Aging as war between chemical and biochemical processes: protein methylation and the recognition of age-damaged proteins for repair.* Ageing Res Rev, 2003. **2**(3): p. 263-85.
63. Lindner, H. and W. Helliger, *Age-dependent deamidation of asparagine residues in proteins.* Exp Gerontol, 2001. **36**(9): p. 1551-63.
64. Brown, M.C. and C.E. Turner, *Paxillin: adapting to change.* Physiol Rev, 2004. **84**(4): p. 1315-39.
65. Cleveland, D.W., *NuMA: a protein involved in nuclear structure, spindle assembly, and nuclear re-formation.* Trends Cell Biol, 1995. **5**(2): p. 60-4.
66. Fant, X., A. Merdes, and L. Haren, *Cell and molecular biology of spindle poles and NuMA.* Int Rev Cytol, 2004. **238**: p. 1-57.
67. Stumpf, M., et al., *The mediator complex functions as a coactivator for GATA-1 in erythropoiesis via subunit Med1/TRAP220.* Proc Natl Acad Sci U S A, 2006. **103**(49): p. 18504-9.
68. Minard, M.E., et al., *The role of the guanine nucleotide exchange factor Tiam1 in cellular migration, invasion, adhesion and tumor progression.* Breast Cancer Res Treat, 2004. **84**(1): p. 21-32.
69. Shi, B.J., et al., *Novel mutations of the ATP2A2 gene in two families with Darier's disease.* Arch Dermatol Res, 2009. **301**(1): p. 27-30.
70. Chao, H.K., K.J. Hsiao, and T.S. Su, *A silent mutation induces exon skipping in the phenylalanine hydroxylase gene in phenylketonuria.* Hum Genet, 2001. **108**(1): p. 14-9.
71. Kimchi-Sarfaty, C., et al., *A "silent" polymorphism in the MDR1 gene changes substrate specificity.* Science, 2007. **315**(5811): p. 525-8.
72. Zhu, H., et al., *A known functional polymorphism (Ile120Val) of the human PCMT1 gene and risk of spina bifida.* Mol Genet Metab, 2006. **87**(1): p. 66-70.
73. Grace, E.A. and J. Busciglio, *Aberrant activation of focal adhesion proteins mediates fibrillar amyloid beta-induced neuronal dystrophy.* J Neurosci, 2003. **23**(2): p. 493-502.
74. Ferhat, L., et al., *The nuclear/mitotic apparatus protein NuMA is a component of the somatodendritic microtubule arrays of the neuron.* J Neurocytol, 1998. **27**(12): p. 887-99.

75. Jiang, Y.M., et al., *Gene expressions specifically detected in motor neurons (dynactin 1, early growth response 3, acetyl-CoA transporter, death receptor 5, and cyclin C) differentially correlate to pathologic markers in sporadic amyotrophic lateral sclerosis.* J Neuropathol Exp Neurol, 2007. **66**(7): p. 617-27.
76. Boise, L.H., et al., *Bcl-2 and Bcl-2-related proteins in apoptosis regulation.* Curr Top Microbiol Immunol, 1995. **200**: p. 107-21.
77. Mendoza-Naranjo, A., C. Gonzalez-Billault, and R.B. Maccioni, *Abeta1-42 stimulates actin polymerization in hippocampal neurons through Rac1 and Cdc42 Rho GTPases.* J Cell Sci, 2007. **120**(Pt 2): p. 279-88.
78. Jacobsen, N.J., et al., *Exclusion of the Darier's disease gene, ATP2A2, as a common susceptibility gene for bipolar disorder.* Mol Psychiatry, 2001. **6**(1): p. 92-7.