University of Pavia
Department of Molecular Medicine

IRCCS Mondino Foundation

PhD course in Translational Medicine

XXXIII cycle

PhD thesis on

Dissecting the genetic bases of Amyotrophic lateral sclerosis and late-onset Pompe disease through Next-generation sequencing

Tutor:
Prof.ssa *Antonella Minelli*
Prof.ssa *Cristina Cereda*

Candidate:
Dr. *Ilaria Palmieri*

Academic year 2019-2020

Index

Abstract

Amyotrophic lateral sclerosis (ALS) is a neuromuscular disorder that affects motor neurons leading to respiratory failure and death in 2 – 4 years after diagnosis. ALS aetiology can be either sporadic (90% of cases) or familial (10% of cases), the latter usually defined by a genetic cause. Genetically, the ALS landscape continues to shift as the number of genes associated with the disease keeps on growing. Moreover, ALS clinical manifestation is complicated by both genetic and phenotypic overlapping with other diseases, such as Frontotemporal dementia (FTD), Hereditary spastic paraplegia (HSP) and Parkinson's disease (PD), but also with other neuromuscular diseases. Also, ALS has now considered an oligogenic disease in which the sum of several variants with small effects in different genes may lead to disease onset. Late-onset Pompe disease (LOPD) is a rare monogenic autosomal recessive disorder, caused by mutations in the *GAA* gene. The molecular genetic testing for LOPD is based on Sanger sequencing analysis of the *GAA* gene, however large deletions/duplications, not detectable through direct sequencing, are usually poorly investigated. LOPD is characterized by highly variable clinical presentation, despite identical or similar genotypes, hence genetic factors other than pathogenic *GAA* mutations may play a role in determining it.

The aims of the two Ph.D. projects were 1) the study of the genetic architecture of 248 sporadic or apparently sporadic ALS patients through Next generation sequencing (NGS), considering pathogenic variants in genes associated with ALS, rare variants that confer susceptibility to the disease and risk factor variants either in ALS genes but also in genes linked to FTD, HSP, PD and other NDs. Specifically, we screened 97 genes in 60 patients (Panel 1), 173 genes in 107 patients (Panel 2), and 277 genes in 81 patients (Panel 3); the investigation of the correlation of known ALS-implicated with our ALS cohort; the study of oligogenic bases for ALS onset in our cohort of patients; 2) The genetic investigation through MLPA of LOPD patients which carried only one pathogenic mutation in the *GAA* gene; the study of the impact of three single nucleotide polymorphisms (SNPs) within the *GAA* gene on protein stability; the identification of LOPD genetic modifiers through the WES analysis on 30 LOPD patients, focusing on 302 genes belonging to six pathways relevant for LOPD. All the genetic analyses have been performed upon genomic DNA extraction. NGS has been done on 248 ALS patients using the SureSelect^{QXT} technology (Agilent Technologies), while NGS has been done on 30 LOPD patients using the

SureSelect^{XT} HumanAllExon V5 + UTRs (Agilent Technologies) capture library. The eVAI (enGenome) web platform has been used for variant annotation. Through NGS, we identified 38 pathogenic variants and 31 likely pathogenic variants (16%) out of 423 variants retrieved in the ALS cohort of patients. This higher percentage compared to the percentage of causative variants retrieved in sporadic patients in the literature, suggests that the enlargement of the number of genes analyzed may be helpful in explaining a larger amount of ALS cases. Among the pathogenic and likely pathogenic variants, 25 were in ALS causative genes, followed by 5 variants in the HSP and PD associated genes respectively and the remaining three variants were in genes linked to other NDs. We established a high degree of genetic heterogeneity among our patients, with 26 different ALS-implicated variants from 14 ALS genes identified among 54,4% of cases. Among the 26 variants, two were significantly associated with our cohort of ALS patients, an in-frame deletion (p.Gly173_Gly174del) within *FUS* and a missense variant (p.Arg208Trp) within *SIGMAR1*. Additionally, 7 patients (3%) of our cohort presented more than one of the listed ALS-implicated variants, suggesting oligogenic bases for these patients. Through MLPA analysis we identified two heterozygous deletions in 2 LOPD patients, one in exon 17 and one in exon 1 of the *GAA* gene. Through NGS we found a total of 176 variants in 103 (34%) out of 302 genes analysed across six virtual gene panels. Specifically, we found 82 variants in 46 genes of the Lysosome pathway, 31 variants in 15 genes linked to the Regulation of autophagy pathway, 45 variants in 28 genes related to the Apoptosis pathway, 5 variants in 5 genes of the Renin-angiotensin pathway, 7 variants in 5 genes of the Glycogen metabolism pathway and 6 variants within 4 genes of the ACE inhibitor pathway. Through statistics analysis, 73 out of 176 (43%) resulted significantly associated with our cohort of patients, however we lack to found genetic modifers for the clinical presentation of the LOPD pathology. This experimental thesis has reaffirmed the heterogeneity of the genetic of ALS, which takes into consideration not only ALS-causative genes and risk factors in genes notably associated with ALS, but consider also genes associated with other NDs such as FTD, HSP and PD, as well as gene related to other motor neuron diseases, highlighting the importance of genetic comorbidity studies. Furthermore, here we confirmed the new vision of ALS as a polygenic disease characterized by oligogenic inheritance. Additionally, here we underline also the importance in the utilization of NGS technology in genetically less complex diseases, such as LOPD, to

find genetic variants associated with the disease that may be at the basis of the clinical variability observed among LOPD patients. Eventually MLPA technology should be used when only one pathogenic mutation is identified.

Introduction

1. Amyotrophic lateral sclerosis

Amyotrophic lateral sclerosis (ALS), also known as Charcot's disease or Lou Gehrig's disease, is a progressive neurodegenerative disorder (ND), characterized by the degeneration of motor neurons in brain and spinal cord.[1] The first description of ALS dates back to the middle '800 when Jean-Martin Charcot used the definition "*amyotrophic lateral sclerosis*" in relation to a progressive muscular atrophy with underlying neurological problems.[2] Onset of ALS usually occurs in late middle life with a progressive muscle atrophy and weakness, that ultimately affects respiratory muscles leading to respiratory failure and death 2–4 years after disease onset.[3] The main neuropathological features of ALS are i) the massive loss of lower motor neurons (LMNs) that project from the spinal cord and brainstem into skeletal muscles, ii) the degeneration and loss of pyramidal cell neurons in the primary motor cortex and of upper motor neurons (UMNs) that project from the motor cortex into the brainstem and spinal cord and iii) the reactive gliosis, such as the hypertrophy of glial cells in the motor cortex and spinal cord in the areas where degeneration occurs.[4]

1.1 Epidemiology

ALS is the most common form of motor neuron disease, with a worldwide prevalence of approximately 6 cases per 100,000 individuals and an annual incidence of 1-2.6 cases per 100,000 individuals.[5] Among Europeans, the ALS mean prevalence is 5.40/100,000 and the mean incidence is 2.8 cases per 100,000 people[4] and affects 1·2–1·5 men for every woman.[6,7] Even if in the majority of cases the disease onset occurs during late adulthood (55-70 years of age), juvenile (before 25 years of age) and "young-onset" (before 45 years of age) ALS cases exist, representing about 1% and 10% of all cases respectively.[4] As for other NDs, such as Parkinson's disease (PD) and Alzheimer's disease (AD), individuals with a family history of ALS tends to have an earlier age of onset compared to individuals without any ALS cases within the family. This may be due to closer surveillance of ALS symptoms among relatives of ALS patients, leading to earlier clinical diagnosis and to the possibility that familial cases have a higher genetic load.[7,8]

1.3 Neuropathology

- **Macroscopic:** most ALS brains do not present gross abnormalities and appear unremarkable, but atrophy of the frontal or temporal cortex and of the precentral-gyrus may occur.[9] At the level of the spinal cord, overall atrophy is observed, and the comparison of the thickness of anterior and posterior nerve roots is a helpful tool to assess the presence and the degree of the anterior nerve root atrophy due to motor neuron loss.[9–11]
- **Microscopic:** microscopic changes include both neuronal and axon loss. There is loss of myelinated axons in the lateral and anterior columns of the spinal cord and decreases in size of anterior horn of the spinal cord; there is degeneration and loss of the large motor neurons in the anterior horn of the spinal cord, lower cranial motor nuclei of the brainstem, and Betz cells in the motor cortex; there is a global reduction of all neurons in the anterior horn, not just the large alpha-motorneurons;[12] there is evidence of reduction in neuron size as well as loss and atrophy of nerve fibers. Additionally, vacuolization (large empty spaces near neurons) and spongiosis (microscopic holes resulting in a sponge like appearance) may occur. Bunina bodies, which are small eosinophilic inclusions in the cytoplasm of motor neurons of the spinal cord and brain stem of ALS, are often detected.[13,14] Bunina bodies are negative for a variety of proteins commonly associated with neurodegeneration, including tau, a-synuclein, and amyloid precursor protein.[15–17] It remains controversial whether or not Bunina bodies are positive for ubiquitin.[18] Their biological significance is unknown. Reactive gliosis in the site of motor neurons degeneration has also been seen,[19] as well as microglia activation.[20]
- **Molecular pathology:** the discovery of ubiquitin-positive round structures firstly in the cytoplasm of anterior horn cells, and lately in neurons of the frontal cortex, temporal cortex, hippocampus, and striatum, in both fALS and sALS patients, lead to a fundamental progress in the understanding of ALS pathology. Importantly, these inclusions were lately found also in FTD patients, signing one of the first link between these two diseases. These ubiquitin-positive inclusions were negative for the presence of other known proteins involved in neurodegeneration, such as tau and α-synuclein, but instead the main component was found to be TDP-43, particularly, the phosphorylated form.[21,22] TDP-43 is a nuclear protein, which, under pathologic conditions, forms cytoplasmic and, in

some cases, nuclear inclusions glial cells (22% of cases), in both neuronal and glial cells (59% of cases), and in neurons only (7% of cases) of both fALS and sALS patients.[23] Strikingly, TDP-43 aggregation is evident in ~97% of all ALS cases.[24] These TDP-43 inclusions are evident in both demented and non-demented patients with ALS, and increase in density with disease evolution, particularly the development of cognitive impairment.[25] The extent of TDP-43 pathology differentiates ALS- FTD from ALS without FTD,[26] and the presence of TDP-43 pathology in extra-motor areas was associated with cognitive impairment in ALS[27] TDP-43 pathology in the orbitofrontal, dorsolateral prefrontal, medial prefrontal cortices and ventral anterior cingulate were associated with executive dysfunction. Language dysfunction was associated with TDP-43 pathology in the inferior frontal gyrus, transverse temporal area, middle and inferior temporal gyri, as well as the angular gyrus. Verbal fluency dysfunction was associated with TDP-43 pathology in the prefrontal cortex, inferior frontal gyrus, ventral anterior cingulate and transverse temporal area. Behavioural abnormalities, however, were associated with TDP-43 pathology in the orbitofrontal and prefrontal cortices as well as the ventral anterior cingulate.[27]

1.2 Genetics of ALS

ALS is characterized by a complex genetic involvement, with a Mendelian pattern of inheritance in some cases but with no discernible family history in the rest. Recently, ALS has been considered as a polygenic disorder due to the contribution to several variants with minor to moderate effects. An inheritance pattern has been observed in the 10% of cases that are considered familial (fALS), while the remaining 90% is considered sporadic (sALS), with no prior family history.[28] In these last cases, heritability may be hidden by small pedigrees, causing familial cases to appear sporadic.[29] Interestingly, about 10% of patients with sALS have mutations in genes that are known to be associated with fALS, leading the dichotomization of ALS into familial and sporadic diseases to be an over-simplification.[30] This genetic overlapping between fALS and sALS may explain similarities in their clinical and pathological features, which suggests a convergence of the cellular and molecular events leading to motor neuron degeneration.[31] Also, the 90% of apparently sporadic cases and incomplete penetrance of several genes linked to familial cases suggest that at least some forms of ALS arise from the interplay of multiple genes, poorly understood

developmental, environmental, and age-related factors, as well as stochastic events.[32] Additionally, beyond genetic founders, also the role of *de novo* mutations has been taken into consideration in the ALS pathogenesis context. However, even if the results are in not clear and in contrast, to date it appears that no differences in the rate of the *de novo* mutations in ALS cases is observed with respect the general population, suggesting that *de novo* mutations are a pathogenic mechanism only in a minority of cases.[33]

Since the discovery of the first gene linked to ALS, the Superoxide dismutase 1 (*SOD1*) in 1993[34], over 130 genes have been associated with ALS, although only a subset have strong and replicated support.[35] To date, 25 genetic loci have been linked to ALS (Figure 1 and Table 1). Most of these genes are autosomal dominant (AD) (i.e.*TARDBP*, *FUS* and *OPTN*), others display an autosomal recessive (AR) inheritance (Spatacsin (*SPG11*), Alsin (*ALS2*) and sigma non-opioid intracellular receptor 1 (*SIGMAR1*)), or have X-linked inheritance (*UBQLN2*). Two different patterns of inheritance (AD and AR) have been also observed for *SOD1*.[36]

Evidence suggests an oligogenic inheritance for ALS, determined by the sum of mild and moderate genes, and genetic pleiotropy, in which a single gene has multiple phenotypic manifestations. Even in the case of these known Mendelian-inherited genes, fALS are often characterized by <50% penetrance and genetic pleiotropy, with evidence of oligogenic and polygenic inheritance in individuals with apparently sporadic disease.[37] Genetic variants that modify the clinical phenotype or that enhance susceptibility to ALS are of huge interest, even if the variants themselves do not cause ALS per se.

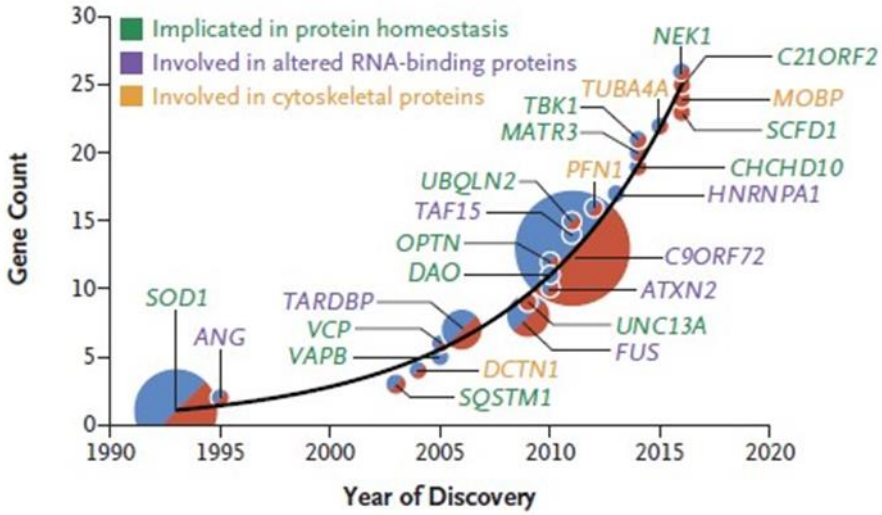

Figure 1| ALS Gene Discovery since 1990. The size of each circle reflects the proportion of all fALS cases associated with that gene. Blue circles indicate genes associated only with fALS, red circles indicate genes associated only with sALS, and circles that are half blue and half red indicate genes associated with both fALS and sALS. ALS genes are clustered in three categories, involving protein homeostasis, RNA homeostasis and trafficking, and cytoskeletal dynamics.[38]

ALS	Gene (RefSeq NM)	Protein (RefSeq NP)	Disease (OMIM)	Inherit.
ALS1	SOD1 (NM_000454.4)	Superoxide dismutase 1 (NP_000445.1)	Amyotrophic lateral sclerosis 1 (MIM:105400)	AD
			Amyotrophic lateral sclerosis 1 (MIM:105400) D90A; D96N	AR
ALS2	ALS2 (NM_020919.3)	Alsin Rho guanine nucleotide exchange factor ALS2 (NP_065970.2)	Amyotrophic lateral sclerosis 2 (MIM:205100)	AR
			Infantile-onset ascending spastic paralysis (MIM:607225)	AR
			Juvenile primary lateral sclerosis (MIM:606353)	AR
ALS3	Unknown			
ALS4	SETX (NM_015046.5)	Senataxin (NP_055861.3)	Amyotrophic lateral sclerosis 4, juvenile ALS (MIM:602433)	AD
			Spinocerebellar ataxia, 1 (MIM:606002)	AR
ALS5	SPG11 (NM_025137.3)	Spatacsin (NP_079413.3)	Amyotrophic lateral sclerosis 5, juvenile (MIM:602099)	AR
			Charcot-Marie-Tooth disease, axonal type 2X (MIM:616668)	AR
			Spastic paraplegia 11 (MIM:604360)	AR
ALS6	FUS (NM_004960.3)	FUS RNA-binding protein (NP_004951.1)	Amyotrophic lateral sclerosis 6, +/- FTD (MIM:608030)	AD
			Tremor, hereditary essential 4 (MIM:614782)	AD
ALS7	Unknown			
ALS8	VAPB (NM_004738.4)	Vesicle-associated membrane protein-B/C (NP_009057.1)	Amyotrophic lateral sclerosis 8 (MIM:608627)	AD
			Spinal muscular atrophy, late-onset (MIM:182980)	AD
ALS9	ANG (NM_001145.4)	Angiogenin (NP_001136.1)	Amyotrophic lateral sclerosis 9 (MIM:611895)	AD
ALS10	TARDBP (NM_007375.3)	TAR DNA-binding protein 43 (NP_031401.1)	Amyotrophic lateral sclerosis 10, +/- FTD (MIM:612069)	AD

ALS11	FIG4 (NM_014845.5)	Polyphosphoinositide 5-phosphatase (NP_055660.1)	Amyotrophic lateral sclerosis 11 (MIM:612577)	AD
			Charcot-Marie-Tooth disease 4J (MIM:611228)	AR
			Polymicrogyria, bilateral temporooccipital (MIM:612691)	AR
ALS12	OPTN (NM_021980.4)	Optineurin (NP_068815.2)	Amyotrophic lateral sclerosis 12 (MIM:613435)	AD/AR
			Glaucoma 1, open-angle, E (MIM:137760)	AD
ALS13	Unknown			
ALS14 FTDALS6	VCP (NM_007126.3)	Valosin containing protein (NP_009057.1)	Amyotrophic lateral sclerosis 14, +/- FTD (MIM:613954)	AD
			Charcot-Marie-Tooth disease, type 2Y (MIM:616687)	AD
			Inclusion body myopathy with Paget disease +/- FTD (MIM:167320)	AD
ALS15	UBQLN2 (NM_013444.3)	Ubiquilin-2 (NP_038472.2)	Amyotrophic lateral sclerosis 15, +/- FTD (MIM:300857)	XLD
ALS16	SIGMAR1 (NM_005866.2)	Sigma non-opioid intracellular receptor 1 (NP_005857.1)	Amyotrophic lateral sclerosis 16, juvenile (MIM:614373)	AR
			Distal spinal muscular atrophy, 2 (MIM:605726)	AR
ALS17 FTDALS7	CHMP2B (NM_014043.3)	Charged multivesicular body protein 2b (NP_054762.2)	Amyotrophic lateral sclerosis 17 (MIM:614696)	AD
			FTD chromosome 3-linked (MIM:600795)	AD
ALS18	PFN1 (NM_005022)	Profilin-1 (NP_005013)	Amyotrophic lateral sclerosis 18 (MIM:614808)	AD
ALS19	ERBB4 (NM_005235.2)	erb-b2 receptor tyrosine kinase 4 (NP_005226.1)	Amyotrophic lateral sclerosis 19 (MIM:615515)	AD
ALS20	HNRNPA1 (NM_002136)	Heterogeneous nuclear ribonucleoprotein A1 (NP_002127)	Amyotrophic lateral sclerosis 20 (MIM:615426)	AD
			Inclusion body myopathy with Paget disease +/- FTD3 (MIM:615424)	AD
ALS21	MATR3 (NM_199189.2)	Matrin-3 (NP_954659.1)	Amyotrophic lateral sclerosis 21 (MIM:606070)	AD
ALS22	TUBA4A (NM_006000.3)	Tubulin alpha-4A (NP_005991.1)	Amyotrophic lateral sclerosis 22, +/- FTD (MIM:616208)	AD
ALS23	ANXA11 (NM_145868.2)	annexin A11 (NP_665875.1)	Amyotrophic lateral sclerosis 23 (MIM:617839)	AD
ALS24	NEK1 (NM_001199397.1)	NIMA related kinase 1 (NP_001186326.1)	Short-rib thoracic dysplasia 6 +/- polydactyly (MIM:263520)	AR

ALS25	KIF5A (NM_004984.4)	Kinesin family member 5a (NP_004975.2)	Amyotrophic lateral sclerosis 22 (MIM:617921)	AD
			Myoclonus, intractable, neonatal (MIM:617235)	AD
			Spastic paraplegia 10, autosomal dominant (MIM:604187)	AD
FTDALS1	C9ORF72 (NM_018325)	chromosome 9 open reading frame 72 (NP_060795)	FTD +/- Amyotrophic lateral sclerosis 1 (MIM:105550)	AD
FTDALS2	CHCHD10 (NM_001301339)	coiled-coil-helix-coiled-coil-helix domain containing 10 (NP_001288268)	FTD +/- Amyotrophic lateral sclerosis 2 (MIM:615911)	AD
FTDALS3	SQSTM1 (NM_003900)	Sequestosome 1 (NP_003891)	FTD +/- Amyotrophic lateral sclerosis 3 (MIM:616437)	AD
			Myopathy, distal, with rimmed vacuoles (MIM:617158)	AD
			Neurodegeneration with ataxia, dystonia, and gaze palsy, childhood-onset (MIM:617145)	AR
			Paget disease of bone 3 (MIM:167250)	AD
FTDALS4	TBK1 (NM_013254)	TANK binding kinase 1 (NP_037386)	FTD +/- Amyotrophic lateral sclerosis 4 (MIM:616439)	AD
			{Encephalopathy, acute, infection-induced (herpes-specific), susceptibility to, 8} (MIM:617900)	AD
FTDALS5	CCNF (NM_001761)	cyclin F (NP_001752)	FTD +/- Amyotrophic lateral sclerosis 5 (MIM: 619141)	-
FTDALS8	CYLD (NM_015247)	CYLD lysine 63 deubiquitinase (NP_056062)	?FTD +/- Amyotrophic lateral sclerosis 8 (MIM: 619132)	-

Table 1| Table of the main genes involved in ALS. For each gene, gene name and Refseq transcript ID, protein name and the Refseq ID, associated diseases (OMIM), and pattern inheritance have been described. **FTD:** Frontotemporal dementia; **MIM**: Mendelian Inheritance in Man; **AD**: autosomal dominant; **AR**, autosomal recessive; **XLD**: X-linked dominant; **+/-:** with or without.

Besides the 25 ALS loci and 8 FTDALS loci, other genes have been associated with ALS TANK binding kinase 1 (*TBK1*), Dynactin Subunit 1 (*DCTN1*), D-amino-acid oxidase (DAO), and Neurofilament Heavy Polypeptide (*NEFH*).

Below, a description of the genetics relative to the principal genes associated with ALS. Genes described are ordered based on the belonging *locus*.

1.2.1 ALS1 – *SOD1*

SOD1 gene has been the first gene associated with ALS, in 1993[34]. It is located on chromosome 21q22.1 and comprises 5 exons spanning 9kb. Mutations in this gene have been linked to ALS 1 (MIM:105400).[39]

SOD1 and encodes for a metalloenzyme of 153 highly conserved amino acids which is one of the two isozymes responsible for destroying free superoxide radicals in the body.[3,39] The function of SOD1 is indeed to protect cells from reactive oxygen species (ROS) by converting them into oxygen and hydrogen peroxide.[40] It is mainly distributed in the cytoplasm; however, it has been found also in the nucleus, lysosomes, and mitochondria (intermembrane space).[41] According to a recent meta-analysis, pathogenic variants in *SOD1* account for approximately 15–30% of fALS and about 2% of sALS cases and it is the second most frequent gene involved in ALS in the European population after *C9ORF72*.[28] Pathogenic mutations are classified into two groups based on their position in the protein: mutations in the beta-barrel region and mutations that affect metal-binding sites.[42] The first ones can result in protein instability, while the second ones alter the catalytic site, resulting in insufficient degradation of ROS with detrimental effects on distal axons of motor neurons. In both cases, *SOD1* mutations destabilize the structure of the protein, leading to oligomerization and formation of aggregates, and affect the normal functioning of the transport from the endoplasmic reticulum (ER) to the Golgi system, causing the destabilization of microtubules.[43,44] To date, over 185 disease-associated variations in *SOD1* have been identified, distributed throughout the gene, giving very diverse clinical phenotypes, severity and disease duration.[45] Rapid disease progression and shorter survival times have been seen in patients with the A4V, H43R, L84V, G85R N86S, and G93A variants, whereas patients with the G93C, D90A, or H46R variants generally have longer life expectancies.[3]

The D90A (aspartate to alanine) mutation is the most common mutation in Europe, accounting for up to 50% of all ALS cases in Sweden and Finland.[46] This variant has been observed both in AD and AR inheritance patterns, even if in the majority of cases it is recessive.[39] Homozygous patients generally exhibit a slowly progressive paresis that starts in the lower limbs and gradually spreads upward, as well as some autonomic dysfunction such as bladder disturbance. By contrast, heterozygous patients display a more heterogeneous phenotype including bulbar with upper limb onset and lower limb onset with a faster progression.[3] The A4V (alanine to valine) is the most common SOD1 mutation in the U.S. population.[47] Clinically, patients with the A4V mutation causes have a sudden onset of symptoms, a rapid disease progression, and LMN dysfunction.[3,39]

1.2.2 ALS6 - FUS

FUS is located on chromosome 16p11.2 and comprises 15 coding exons spanning 12kb. Mutations in this gene have been linked to ALS 6 with or without FTD (MIM:608030). *FUS* encodes a multifunctional ubiquitously expressed protein, which is part of the heterogeneous nuclear ribonucleoprotein (hnRNP) complex.[48] Capable of binding to both single- and double-stranded DNA, as well as to RNA, FUS is implicated in cellular processes that include regulation of gene expression, maintenance of genomic integrity, and mRNA/microRNA processing.[48] FUS shuttles between the nucleus and cytoplasm and its accumulation in the cytoplasm is strongly related to an aberrant stress granules (SGs) formation which play a role in protein aggregation both in ALS and Frontotemporal dementia (FTD).[49]

Over 50 AD *FUS* variants have been identified in ALS patients, with a mutation frequency of approximately 5% in fALS and less than 1% in sALS patients. *FUS* mutations have been reported also in FTD patients.[50,51] The great majority of ALS-causing *FUS* mutations impairs its nuclear import. As a result, FUS protein accumulates in cytoplasmic inclusions, affecting RNA correct transcription and translation and DNA damage repair which are known mechanisms of neurodegeneration. The majority of *FUS* mutations are missense, even if deletions, insertions, nonsense, and splicing mutations have been also reported.[52] More than 60% of cases with *FUS* mutations show disease onset before 45 years of age, with many juvenile ALS cases presenting with disease onset around age twenty.[53] Patients harbouring *FUS* mutations had various onset sites (in the arms, legs, or bulbar muscles), young-onset and a rapid course for most of the patients.[54,55]

1.2.3 ALS10 - TARDBP

TARDBP is located on chromosome 1p36.22 and comprises 7 exons spanning 13kb. Mutations in this gene have been linked to ALS 10 with or without FTD (MIM:612069).[56,57] *TARDBP* encodes for the TDP-43 protein, which is a highly conserved heterogeneous nuclear ribonucleoprotein (hnRNP), made of 414 amino acids, involved in RNA metabolism (regulation of transcription, pre-RNA splicing, translation, mRNA stabilization, and microRNAs processing). TDP-43 resides mostly in the nucleus, but up to 10% of TDP-43 is present in the cytoplasm. Under pathophysiological conditions, TDP-43 is cleared from the nucleus and accumulates in the cytoplasm, as a consequence of an impairment of nuclear transport and

solubility.[3,58] This accumulation leads to the formation of ubiquitin-positive cytoplasmic neuronal inclusions in the brain and spinal cord, hallmarks of ALS.[59]

To date, several mutations have been identified in this gene, affecting both fALS and sALS, with a prevalence of ~3% and ~1.5% cases respectively.[60] The majority of variants are missense mutations located in a mutational hotspot in the sixth coding exon of the *TARDBP* gene, thus affecting the C-terminal portion of the protein and ultimately leading to protein instability, cytoplasmic mislocalization, altered interaction with other proteins, and to an increased propensity to aggregate.[39] The A382T and the G295S are the commonest missense mutations in TARDBP.[60] A382T mutation was identified as a founder mutation in ALS patients from Sardinia, with a mutational frequency of about 30%.[61] This mutation was also found in the homozygous state in two sALS cases, who displayed hypokinetic parkinsonian syndrome and FTD.[62] Interestingly, most of the *TARDBP* mutations identified are carried by patients with "classical" ALS phenotype - in terms of age and site of onset of the disease – with no cognitive defects. However, as in the case described above, clinical heterogeneity with signs of dementia and parkinsonism has been demonstrated. Other *TARDBP* mutations in ALS patients that lead to cognitive decline have been observed: the K263E mutation, identified in a patient who developed FTD and supranuclear palsy[63], as well as the N267S mutation, which was found in an FTD patient[64] and in a patient with idiopathic PD.[65]

1.2.4 ALS11- FIG4

FIG4 (FIG4 phosphoinositide 5-phosphatase) gene is located on chromosome 6q21 and comprises 23 coding exons spanning 134kb.[66] Mutations in this gene have been linked to ALS 11 (MIM:612577).[67] *FIG4* encodes a phosphoinositide 5-phosphatase of 907 amino acids, which belongs to the SAC domain-containing protein family and regulates the phosphatidylinositol-3,5-bisphosphate PI(3,5)P2, an intracellular signaling lipid with a key role in endosomal vesicle trafficking.[68] Indeed, FIG4 localizes on the cytoplasmic surface of endolysosomal vesicles.[69]

Homozygous mutations in *FIG4* cause the recessively inherited Charcot-Marie-Tooth neuropathy type 4J (CMT4J), a severe form of CMT, with early-onset and involvement of both sensory and motor neurons. However, in 2009, heterozygous mutations in *FIG4* have been implicated in both adult-onset sALS and fALS, with a frequency of 1-2%.[66] Indeed, Chow *et al.* identified two missense (D53Y and R388G),

two protein-truncation (R183X and Q403X) and two splicing (c.1386 5G>T and c.67-1G>T) mutations, all resulting in a significant or complete loss of FIG4 phosphatase activity.[70] Osmanovic and colleagues found the F254Sfs*8 frameshift variant, suggesting that *FIG4* is sensitive to loss of function mutations.[68] In mouse models Fig4, loss of function was found to cause neuronal degeneration in the central nervous system (CNS) including spinal motor neurons and peripheral neuropathy.[71] However, the contribution of *FIG4* variants to ALS pathogenesis has been still under debate. Clinically, patients with *FIG4* mutations have a predominance of UMN signs and show a disease duration longer (mean 9 years) compare to non *FIG4* variant carriers[68], but patients with LMN signs and fast progression have been observed too.[70]

1.2.5 ALS12 - OPTN

OPTN is located on chromosome 10p13 and comprises 16 coding exons spanning 38kb, arisen by gene duplication at the time of early vertebrates from NF-kB essential modulator (NEMO), the regulatory component of the NF-kB pathway.[72,73] Mutations in this gene have been linked also to ALS 12 (MIM:613435).[73,74] Optineurin is a 577 amino acids ubiquitin-binding protein, highly conserved among species, which functions as a regulatory/adaptor protein in multiple cellular processes, including signal transduction, protein degradation, trafficking, and other cellular processes, commonly regulated by protein conjugation to ubiquitin. At the Golgi, optineurin regulates post-Golgi transport and exocytosis; whereas, in the nucleus, it is involved in the regulation of cell division and gene transcription.[72] Optineurin is ubiquitously expressed; however, a higher expression has been detected in the brain, retinal ganglion cells, heart, skeletal muscle, and spleen.[72] In 2010, the homozygous deletion of exon 5 (Δex5), the homozygous Q398X mutation, and the heterozygous E478G mutation, were identified in Japanese ALS patients, being the first link between ALS and *OPTN*. To date, about 61 putative optineurin mutations have been described in ALS patients, including 36 missense mutations, three delete mutations, eight nonsense mutations, nine frameshift mutations, and five splicing mutations.[75] *OPTN* variants have been identified either in fALS and sALS patients and seem to be more frequent in Japanese and Chinese patients, accounting for 1-4% of fALS and less than 1% of sALS cases, whereas they comprise <1% in the Caucasian population.[75] The nature, frequency, and penetrance of optineurin mutations vary

greatly with ethnicity: indeed, only a few mutations are shared between different populations, suggesting a genetic founder effect.[72] Except for a few mutations, found only in the homozygous state, the majority of the *OPTN* variants found in ALS cases are heterozygous, suggesting that haploinsufficiency is sufficient to cause or contribute to ALS.[72] The clinical phenotype in *OPTN* ALS carriers is not homogeneous: some individuals show relatively slow progression and long duration,[76,77] whereas other carriers display aggressive progression and short survival.[78,79]

1.2.6 ALS14 - VCP

VCP is located on chromosome 9p13.3 and comprises 17 coding exons spanning 17kb. Mutations in this gene have been linked also to ALS 14 with or without FTD (MIM:613954).[80] *VCP* codifies for an ATPase ubiquitously expressed. Thanks to numerous cofactors, VCP functions as a molecular chaperone, linked to several cellular activities including membrane dynamics, protein quality control, DNA damage response, apoptosis, protein degradation and autophagy. However, the underlying molecular mechanism in the disease is not fully understood yet.[81] Missense mutations in the *VCP* gene are responsible for various AD-inherited diseases. The first to be associated with *VCP* was the Inclusion body myopathy associated with Paget disease of the bone and FTD (IBMFTD), a late-onset multisystemic disorder, followed by ALS, PD, HSP, and CMT disease type-2 (HMSN2).[81,82] Moreover, some pathogenic variants in VCP have been found for both IBMFTD and ALS[43], representing one of the first examples of comorbidity of ALS with other diseases VCP-linked. To date, more than 50 disease mutations have been identified in the *VCP* gene, accounting for 1%-2% of fALS cases.[83] H155C and H155R are the commonest mutations, with the former being more severe compared to the latter. The key role of VCP protein is the translocation of misfolded proteins from the ER, followed by their subsequent degradation by the proteasome and through autophagy. Mutated VCP is responsible for causing the accumulation of misfolded proteins in the ER, decreasing proteasomal activity and impairing autophagy; moreover, it has also been shown that mutated VCP causes mitochondrial dysfunction and impaired ATP production.[84,85] Clinically, VCP carriers display heterogeneous phenotypes, from the IBMFTD to the classic ALS, to the CMT2Y[86,87] Patients from Johnson et al. displayed

classic ALS with upper and lower motor signs affecting all four limbs and bulbar musculature and mild frontal lobe dysfunction.[88]

1.2.7 FTDALS1 - *C9ORF72*

The *C9ORF72* gene comprises 12 coding exons spanning 27kb, is located on chromosome 9p21 and alterations in this gene have been linked to FTD and/or ALS 1 (MIM:105550).[89] The gene is transcribed into three different transcript isoforms: transcript variants 2 and 3 encode an isoform of 481 amino acid in length (isoform a), whereas transcript variant 1 encodes a shorter protein of 222 amino acid (isoform b) (Figure 2).[90] The three transcription isoforms codify eventually for two protein isoforms, named a and b, whose function is still unclear. Immunohistochemical analysis of the brain showed that C9ORF72 is largely a neuronal cytoplasmic protein and it also seemed to localize at large presynaptic terminals.[91] Although the exact function of C9ORF72 is not known, it has been postulated that C9ORF72 acts as guanine exchange factors for the Rab family of GTPases, involved in the membrane-trafficking, including the autophagic pathway.[92,93]

In 2011, the hexanucleotide repeat expansion (HRE) GGGGCC $(G4C2)_n$ in the non-coding region of the *C9ORF72* gene has been individuated as the cause of the disease both in ALS and FTD patients.[89,94] Three hypotheses have been made to explain how *C9ORF72* HRE leads to both ALS and FTD:[95]

1) The expansion leads to reduced transcription of C9orf72, leading to haploinsufficiency.
2) Repeat-containing sense and antisense RNA transcripts accumulate and sequester RNA binding proteins.
3) Five potentially toxic, repetitive dipeptide proteins (DRPs) originate by the translation of both sense and antisense RNA by repeat-associated, non-ATG initiated (RAN) translation. These DRPs vary in their toxicity through direct binding and sequestration of proteins, such as proteasomal subunits, or by impairing rRNA synthesis, ribosome biogenesis, and translation.[96]

To date, the HRE is considered the major genetic cause of ALS, causing up to 40% of familial cases and about 5%–10% of sporadic cases. Likewise, the same C9ORF72 HRE is present in approximately 5%–10% of all patients with FTD, and in up to 30% of patients with both diseases.[97]

Healthy individuals generally carry between two and twenty-three $(G4C2)_n$ repeats, whereas in patients with ALS the actual size of the HRE varies from hundreds to thousands of repeats.[85] It should to be taken into consideration, however, that the length of the HRE which is usually assessed in blood, may not correlate with those in the brain.[98] Studies of repeat sizes between generations in *C9ORF72* families are almost lacking. One study identified a family with suggestive evidence for genetic anticipation over three generations,[98] however, this phenomenon common to other pathogenic repeat expansions is still under debate in the context of ALS.

Clinically, patients with the *C9ORF72* repeat present with widely variable symptoms on the ALS/FTD spectrum, including ALS, FTD, and ALS-FTD. In ALS, *C9ORF72* expansion have been reported to have a higher incidence of bulbar-onset, cognitive impairment with earlier disease onset, and accelerated progression compared with patients without the expansion.[99]

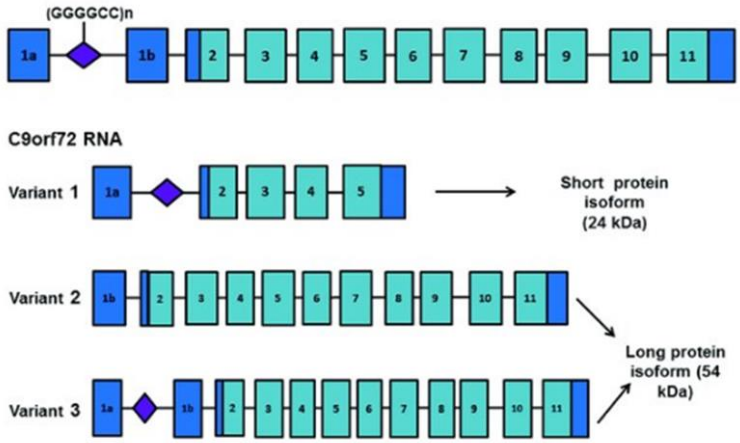

Figure 2| Schematic representation of the C9orf72 gene relative RNA transcript variants. Dark blue boxes represent the 5' and 3' untranslated regions, light blue boxes represent the coding exons, and the GGGGCC HRE is in purple. Differential selection of transcription start and termination sites generates three different RNA transcripts.[90]

1.3 Pathophysiology

Despite the effort made in the last decades, the aetiology of ALS remains poorly understood yet. To date, the only recognized risk factors are age and family history, while a definite environmental risk factor has not been identified.[5] Pathological features of ALS include skeletal muscle atrophy, atrophy of the motor cortex, and pallor and sclerosis of the pyramidal, the corticospinal and corticobulbar tracts, together with thinning of the hypoglossal nerves and the ventral roots of the spinal cord.[37] Microscopic examination typically reveals a depletion of more than 50% of spinal motor neurons and diffuse astrocytic gliosis and microglial infiltration in the

grey and white matter of the spinal cord. Axonal loss, gliosis, and myelin pallor are distinctive features of the corticospinal tracts while astrocytic gliosis is usually observed in the motor cortex, together with a variable depletion of UMNs. Skeletal muscle shows features of denervation and re-innervation.[37] Even though the causative pathogenic mechanism is still not fully understood, several cellular players and molecular processes in ALS have been investigated, highlighting how multiple factors rather than a single initiating event contribute both to the development and progression of the disease (Figure 3).[3,100] The main molecular mechanisms involved in ALS pathogenesis are the following:

- ***RNA metabolism defects:*** RNA processing defects and disruption of RNA binding proteins are increasingly recognized as critical determinants of ALS and other NDs. RNA binding proteins are critical to the maintenance of the transcriptome and they are involved in several aspects of RNA metabolism, including splicing, transcription, transport, translation, and storage in SGs.[101] Many RNA-binding proteins such as FUS[102,103], TDP-43[104,105], heterogeneous nuclear ribonucleoprotein A1 (*hnRNPA1*) and heterogeneous nuclear ribonucleoproteins A2/B1 (*hnRNPA2B1*)[106] (all associated with ALS) are involved in pre-mRNA processing, leading to splicing defects if altered.[105,107] *ELP3* (elongator protein 3) is responsible for RNA translation, and mutations in this gene have been associated with ALS.[108] Also, mutations in the *SETX* (senataxin) gene[109], which regulates the transcription of ribosomal RNA and in the ANG (angiogenin) gene[110], which has a role in RNA processing, have been associated with ALS and might lead to RNA metabolism defects.
- ***Mitochondrial dysfunction:*** Mitochondria are organelles fundamental for ATP generation, calcium buffering, apoptosis regulation and are critical for the oxidative stress balance since they produce the majority of ROS like superoxide radicals.[111] Evidence suggests that the metabolic changes observed in several NDs derive from the disruption in mitochondrial function and the reduction in ATP production.[112,113] Indeed, a decreased electron transport chain efficiency[114,115] and reduced mitochondrial enzyme activity have been verified to occur in ALS animal models as well as in fALS patients.[116,117] In some patients harbouring *SOD1*[118,119] and *C9ORF72*[120] mutations, as well as in SOD1G93A[121] and TDP-43A315T[122] murine models, changes in mitochondrial morphology and defects in respiratory chain function due to oxidative damage to mitochondrial enzymes

have been observed. Also *FUS* mutations have been seen to alter mitochondrial biogenesis.[123,124] Accumulation of cytoplasmatic FUS may alter mitochondrial functions, inducing mitochondrial fragmentation and ROS overproduction.[123,124] Altered Optineurin and TBK1 have been also seen to affect the mitophagic flux, leading to the accumulation of non-functional mitochondria and motor neuron death.[73,125] Moreover, damaged mitochondria and alterations in their functions can also affect calcium homeostasis, increasing neurons' sensitivity to glutamate excitotoxicity and the risk of motor neuron damage.[101]

- ***Oxidative stress:*** as above mentioned, oxidative stress and ROS accumulation have been suggested as a primary initiating factor in ALS pathogenesis.[5] The role of oxidative stress in ALS was highlighted by the discovery of pathogenic variants in the *SOD1* gene, which encodes a major antioxidant defense protein.[126] Oxidative stress interacts with and potentially exacerbates other pathophysiological processes contributing to motor neuron injury, including excitotoxicity, ER stress, protein aggregation, mitochondrial impairment, and alterations in signalling from microglia and astrocytes.[127] ROS generation could result from the activity of NADPH oxidase. The ataxin-2 polyglutamine (PolyQ) protein (encoded by the *ATXN2* gene) can interact with NADPH oxidase, leading to ROS production and accumulation, DNA damage, and mitochondrial distress.[128] *ATXN2* intermediate-length PolyQ expansions (27-33 glutamines) have been found to be significantly associated with increased risk for ALS.[129]
- ***Impaired DNA repair:*** oxidative stress and/or mitochondrial dysfunction throughout the cell life cycle can trigger DNA damage.[101] Hence, a further mechanism suggested to be involved in ALS pathogenesis is the impairment in DNA repair. Mutations in genes that codify for proteins involved in DNA damage repair have been seen to confer susceptibility to ALS, such as SETX[130], NEK1,[131] FUS,[132] and C21orf2 (Cilia And Flagella Associated Protein 410)[133].
- ***Axonal transport disruption***: Motor neurons are known to have exceptionally long axons (up to 1 m in length), thus requiring tight regulation and a high amount of energy for the axonal transport of organelles such as mitochondria or of molecules like proteins, lipids, and RNA toward the synapse and back.[134] Axon maintenance relies mostly on the neurofilaments: neuron-specific intermediate filaments composed of light, medium, and heavy-weighted neurofilaments (NF-L, NF-M, and NF-H, respectively) and α-internexin or peripherin.[135,136] Indeed, the

impairment of this structure has been reported in ALS.[137] Profilin-I, a small actin-binding protein, is involved in G-actin polymerization, and mutations in PFN1 gene have been seen to impair the attachment of actin to the microtubules, probably impacting the anterograde and retrograde transport and leading to UMN and LMN death.[138139] Since the axonal transport also relies on microtubules (made of dimers of α- and β-tubulin) thank either to the dynein-dynactin complex and to kinesins, mutations in TUBA4A[140], DCTN1[141], and KIF5A[142] have been described in ALS cases.

- ***Glutamate excitotoxicity:*** Glutamate excitotoxicity may result from defects in the uptake of glutamate which ultimately leads to an excessive neuronal Ca2+ intake, mitochondrial dysfunction and increased ROS.[143,144] *SOD1* mutations have been implicated in the glutamate excitotoxicity hypothesis, indeed in literature, SOD1 G93A have been described with increased glutamate release mice.[145] *C9ORF72* has also been implicated in the glutamate excitotoxicity hypothesis. *C9ORF72*-expanded patients present high levels of glutamate in the CSF[146] and iPSC motor neurons from ALS patients were found to have pathological accumulations of glutamate receptors.[147,148] Although this evidence, glutamate excitotoxicity as a pathogenic mechanism involved in ALS is still under debate. Despite this, Riluzole, an anti-glutamatergic drug, is the most used treatment of ALS so far.[149,150]
- ***Toxic protein aggregation and impaired protein homeostasis:*** one of the principal neuropathological hallmarks of ALS is the aggregation and accumulation of proteinaceous inclusions in motor neurons. Protein aggregates positive for TDP-43[151], neurofilament[152], FUS[124], or SOD1[153] have been found in almost all ALS patients, with TDP-43 being the most present.[154] Specifically, TDP-43 proteinaceous inclusions are present in the 97% of ALS patients[155], regardless their aetiology, while protein aggregation and inclusion of FUS and SOD1 occur only in patients harbouring *FUS* and/or *SOD1* mutations respectively.[156] Mutations in other ALS genes, such as *OPTN*[73], *UBQLN2*[157], *DNCT1*[158], *VCP*[159], and *MATR3*[160], have been demonstrated to increase the propensity of the respective protein to aggregate, leading directly or indirectly to the impairment of the proteasomal or the autophagic machinery of the cell.[119] *C9ORF72* mutations have been described to interfere with the autophagy pathway, leading to an increase in the number of cytoplasmic inclusions.[161,162]

Dysregulation of the Ubiquitin-Proteasome System (UPS) has been also linked to ALS, upon the finding of ubiquitin-positive protein inclusions and the identification of mutations in genes encoding proteins involved in protein clearance by UPS.[163,164]

- ***Endosomal pathway and vesicle secretion:*** Extracellular vesicles (EVs) comprehend apoptotic vesicles, microvesicles, and exosomes, and their cargo may have an impact on the functionality of the recipient cells.[165,166] EVs are becoming interesting in the ALS context, either for their role in the pathogenesis and as biomarkers.[167,168] Studies demonstrated that EVs are responsible to carry and spread misfolded proteins, such as SOD1, TDP-43 and FUS, responsible for motor neuron death.[169–171] Moreover, it has been suggested that multivesicular bodies accumulated within the cytosol of sporadic muscle cells may disrupt these cells.[172] Affected exosomal trafficking has been observed also in C9orf72 knockdown cell lines[173], C9orf72 knockout mice[174], and in iPSC-derived motor neurons from ALS patients' fibroblasts.[175] VAPB gene which codifies for the vesicle-associated membrane protein-associated protein B has been associated with ALS (VAPB).[176] The VCP gene, strongly associated with ALS, encodes for a protein that interacts with clathrin to form early endosomes as well as with the autophagy pathways.[177] Other genes such as the *FIG4*[178,179] or the SPG11 (spastacsin)[180], all linked to ALS, are associated with the blockade of lysosomal clearance which could potentially lead to vesicle secretion.[176]
- ***Neuroinflammation:*** Increasing evidence shows that neuroinflammation plays a role in ALS pathophysiology. Neuroinflammation associated with neuronal loss is characterized by microglia and astrocyte activation, overproduction of inflammatory cytokines and T lymphocytes infiltration. Microglia detect injuries that disrupt brain homeostasis, responding with morphological changes and the release of cytokines and chemokines to clear pathogens or debris. The inflammatory protein secretion leads to the potentially neurotoxic activation of astrocytes which may contribute to the death of neurons and oligodendrocytes. The role of microglia in ALS pathogenesis was highlighted by the identification of several ALS linked genes that influence the function of, and are highly expressed in microglia, including *C9ORF72*, Progranulin (*PGRN*) and *TBK1.*[3,181]

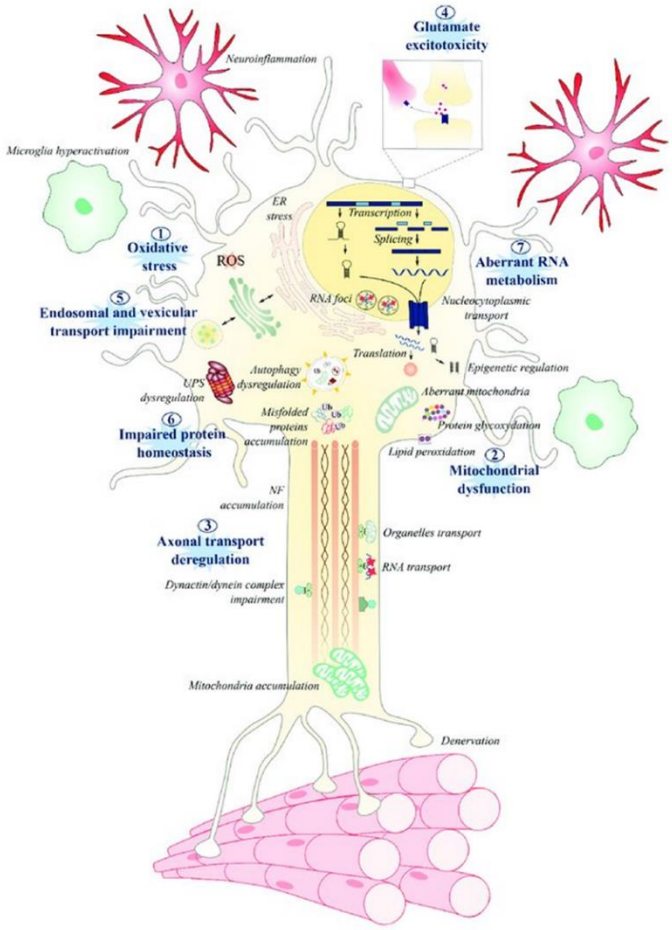

Figure 3| Lists of the pathophysiological mechanisms of ALS.[101]

1.4 ALS Phenotypes

It is well known that ALS clinical spectrum includes different phenotypes depending on the degree of the involvement of spinal and bulbar UMN and LMN. Accordingly, eight distinctive ALS clinical phenotypes have been identified:[182–184]

1. *Classic*: the onset of symptoms occurs in the upper or lower limbs, with definite but not predominant pyramidal signs. This phenotype decrease with age, has the highest incidence among men compared to the other phenotypes, and patients with classic ALS have an intermediate outcome.
2. *Bulbar*: dysarthria and/or dysphagia, tongue wasting, fasciculation and no peripheral spinal involvement are the main clinical features for the first six months after symptoms onset, time by which also pyramidal signs start to become evident. This phenotype increases with age, being the most represented among elder patients, and does not display incidence differences among genders.
3. *Flail arm*: progressive, predominantly proximal, weakness and wasting in the upper limbs characterize patients. Under this category also fit patients with

altered deep tendon reflexes or Hoffman sign in the upper limbs but without hypertonia or clonus. Functional involvement had to be confined to the flail limbs for at least 12 months after the onset of symptoms. This phenotype is rare, more prevalent in men and with a relatively good prognosis.

4. *Flail leg:* progressive distal onset of weakness and wasting in the lower limbs are the main features. Patients with pathological deep tendon reflexes or Babinski sign in the lower limbs but without hypertonia or clonus are also embedded in this category. Patients with proximal weakness with no distal involvement at presentation are classified as classic ALS. This phenotype is the third more common phenotype, has an equal incidence in both genders and a relatively bad outcome.
5. *Pyramidal*: severe spastic para/tetraparesis, associated with one or more of the following signs: Babinski or Hoffmann sign, hyperactive reflexes, clonic jaw jerk, dysarthric speech and pseudobulbar affect, are the main clinical features. Spastic paresis may appear at the beginning or at the late stages of the disease. Clearcut signs of LMN impairment from the onset of the disease are also present, as indicated by muscle weakness and wasting and by the presence of chronic and active denervation at the EMG examination in at least two different sites. It is more common in young-onset with a benign outcome
6. *Respiratory*: respiratory impairment at onset is the prevalent clinical sign. It is defined as orthopnoea or dyspnoea at rest or during efforts, with only mild spinal or bulbar signs in the first 6 months after onset. These patients showed signs of UMN involvement. It is more common in males and, together with the bulbar phenotype, is has the worst prognosis.
7. *Pure lower motor neuron (PLMN):* progressive LMN involvement is evident by electromyography. Patients with motor conduction block(s) on extensive standardized nerve conduction studies, clinical UMN signs, family history for diseases that mimics motor neuron disease are excluded. Neuroimaging tests are also performed to exclude structural lesions. Patients with this phenotype have longer survival compared to patients with any other phenotype besides PUMN (median survival time 7.3 years).
8. *Pure upper motor neuron (PUMN):* severe spastic para/tetraparesis, Babinski or Hoffmann sign, hyperactive reflexes, clonic jaw jerk, dysarthric speech and pseudobulbar affect are present. Patients with signs of LMN involvement,

according to the El Escorial criteria, during the follow-up, and those with a history of disease that mimics motor neuron disease, family history of spastic paraparesis/tetraparesis and mutation of genes related to hereditary spastic paraplegia are excluded. As the pyramidal phenotype, it is more common among younger patients and has a benign outcome.

The stratification of ALS patients according to clinical phenotypes may be useful for improving the design of clinical trials and for more tailored therapies. Moreover, a better knowledge of the factors, such as biochemical and genetics, that underlie the different phenotypic expression of ALS would be important for identifying the molecular mechanisms of the disease.

ALS patients may be also subdivided based on the disease progression rate progression into three main groups: fast, intermediate and slow.[185] Patients with a progression rate less than 0.5 (point/month) were defined as slow (ALS-slow), patients with a progression rate between 0.5 and 1.0 were defined as intermediate (ALS-intermediate), and patients with a progression rate more than 1.0 were defined as fast (ALS-fast).

1.5 Diagnosis and Biomarkers

ALS diagnosis is a challenging process due to its highly heterogenic clinical picture and the comorbidity with other NDs. Since there are no diagnostic tests, ALS diagnosis is mostly clinical, based on the Revised El Escorial World Federation of Neurology Criteria[79] that is based on the following three factors: i) evidence of LMN degeneration on either clinical examination or through electrophysiological changes; ii) evidence of upper UMN impairment on clinical examination and iii) the presence of either of the above in more than one region of the body (Figure 4). Obviously, alongside with the accordance with the El Escorial criteria, other alternative diseases that could explain the combination of progressive upper and lower motor neuron dysfunction must be ruled out. To exclude other diseases Nerve conduction studies and electromyography may be helpful, extending the clinical examination and aiding diagnosis of ALS. Eventually, patients are classified on a spectrum ranging from ‘‘suspected ALS’’ to ‘‘definite ALS’’.

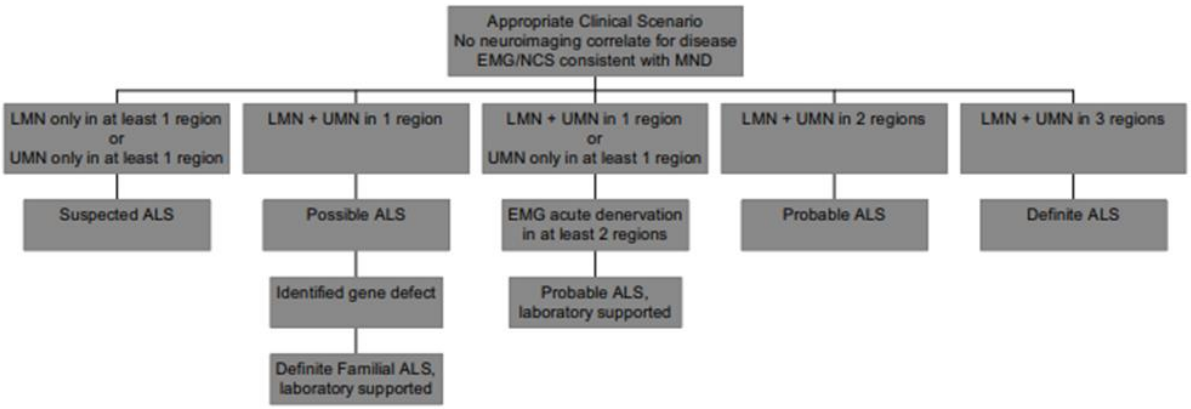

Figure 4| The revised El Escorial criteria for the ALS diagnosis.[186]

However, despite the establishment of clinical criteria and the growing public awareness of ALS, studies from the last three decades have reported a delay of 10–16 months to get an ALS diagnosis after the first arise of symptoms.[187] Here, the urgency to find early biomarkers to shorten this delay which may have a negative impact on therapeutic effectiveness. To date, numerous clinical and molecular biomarkers have been proposed for the diagnosis of ALS[188]. Muscle strength and respiratory function have been extensively studied as markers of disease progression; however, even if they are considered informative, they cannot be used as early markers because they appear only in the advanced stages of the disease[100]. Among the other clinical biomarkers proposed, neurophysiological approaches have been tested to detect motor neuron damage, such as the study of nerve conduction (axonal degeneration), the estimation of the number of motor units through Motor Unit Number Estimation (MUNE) and Motor Unit Number Index (MUNIX), the electrical impedance myography (focused on the functional integrity and muscle structure), and the transcranial magnetic stimulation (functional integrity of neurons)[189,190]. Various imaging techniques have been also proposed as biomarkers and these include single-photon emission computed tomography (SPECT), positron emission tomography (PET), functional magnetic resonance imaging (fMRI), and diffusion tensor imaging (DTI)[190]. Parallelly, also many other molecular biomarkers are under study in these last years. Extracellular vesicles (EVs) are small reservoirs of different molecules important in cell-to-cell communication. As they may have a role in the spreading of misfolded protein between cells, like TDP-43 and SOD1, they have gained much attention in the field of NDs, especially for ALS[191,192]. Another option is represented by circulating RNAs, either coding and non-coding. To date, the most promising miRNAs include miR-206, miR-133b, miR-27a, mi-338-3p, miR-183,

miR-451, let-7 and miR-125b[193,194]. Very recently, also lipids came into attention, with cholesterol, triglycerides and fatty acids being the most investigated. Indeed, it is becoming increasingly accepted that dyslipidemia, should be considered as part of the pathological ALS process[195]. However, despite the wide range of detectable biomarkers, to date, they do not appear to be effective enough for the early stages of the disease but are more useful for monitoring disease progression.

1.6 Existing and future treatments

Currently, the widest treatment used for ALS is Riluzole. This drug has been approved in 1995 and has shown modest benefit to some ALS patients, attenuating disease progression, by inhibiting glutamate release.[3] Since then, several other molecules (anti-inflammatory, antioxidative, anti-glutamatergic, neuroprotective, and neurotrophic compounds) have been investigated as potential treatments for ALS, but the great majority have failed to demonstrate efficacy in human clinical trials. Although success has been limited, there is hope that alternative approaches, including RNA-based therapeutics may be more effective. Two main strategies of RNA-targeted therapeutics have been investigated: short interfering RNA (siRNA) and antisense oligonucleotides (AOs). Several AO drugs have received FDA approval in the last years to treat a variety of conditions, including the treatment of ALS.[3]

Another option could be personalized medicine, a medical scheme that incorporates genetic, clinical diagnostics and environmental information to individualize patient care. Personalized medicine indeed is an important advancement as subtypes of patients may respond differently to potential disease-modifying therapies, so taking complete genetic information into account when assessing clinical trial outcomes may help in determining treatments that are most suitable for particular subgroups of patients.[3]

In the absence of effective pharmacological treatments, symptomatic interventions and supportive care remain the keystone of ALS management, in order to improve the quality of life.[100]

1.7 Frontotemporal dementia (FTD)

FTD (or Pick disease) described for the first time by Arnold Pick in 1892[196] is an insidious neurodegenerative clinical syndrome characterized by progressive atrophy and neural loss in the frontal and temporal cortices. Frontal atrophy subsequently leads to personality and behavioral changes, as well as to gradual impairment of language skills.[85] FTD is the third most common form of dementia, after AD and Lewy bodies Dementia (LBD), and is a leading type of early-onset dementia with an incidence of approximately 4.0 new cases per 100,000 people per year.[197] The average age of onset is 45 – 65 years of age, however, cases younger than age 30 have been reported.[198]

FTD may be divided into three primary subtypes:

1. BV-FTD: behavioral variant, associated with early behavioral and executive deficits. It is the most common form, accounting for up to 60% of FTD cases, and it is the most frequent form associated with motor neuron disorders (MNDs), such as ALS;[199]
2. SV-PPA: semantic variant primary progressive aphasia, a progressive disorder of semantic knowledge and naming;[200]
3. NFV-PPA: non-fluent variant primary progressive aphasia, with progressive deficits in speech, grammar, and word output.[201]

As FTD progresses, symptoms of the three clinical variants can converge into an initial focal degeneration that becomes more diffuse and spreads affecting large regions in the frontal and temporal lobes.[202] Patients gradually develop global cognitive impairment and motor deficits, including parkinsonism, and MND. At late stages, patients display difficulties in moving, eating, and swallowing. Death usually happens about 8 years after symptom onset and is typically caused by pneumonia or other secondary infections.[203] Similar to ALS and other neurological disorders, there is no effective treatment for FTD.[204]

1.7.1 FTD and ALS

It is well known that some ALS patients manifest cognitive decline and behavioral changes. The nature of these peculiar features was associated with FTD at the beginning of the ‘900, through neuropathological studies.[205,206] Since then, ALS and FTD are recognized as part of a broad ND called “The frontotemporal dementia-motor neuron disease continuum” (Figure 5).[207]

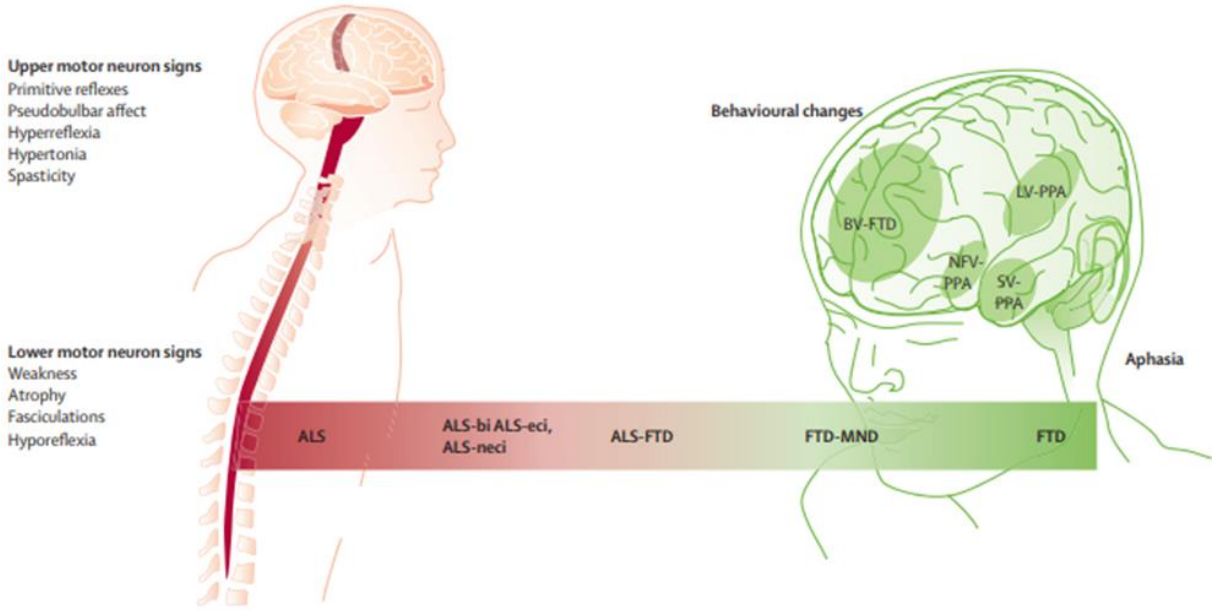

Figure 5| Amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) as part of a spectrum disorder. At the extremities of the spectrum, there are pure ALS and pure FTD. ALS is categorized as ALS-eci, ALS-neci or ALS-bi if executive dysfunction, impairment in other cognitive domains or behavioral changes are present respectively. About 5–10% of patients with ALS also have FTD. Patients with MND–FTD have a primary diagnosis of FTD AND motor neuron involvement develops as the disease progresses but not to full ALS. ALS-eci=ALS with evidence of executive dysfunction. ALS-neci=ALS with no executive dysfunction but impairment in other cognitive domains. ALS-bi=ALS with behavioral changes.[100]

A series of cross-sectional studies have shown that up to 50% of ALS patients develop cognitive and behavioral difficulties matching FTD.[204,208] Comparably, up to 30% of FTD patients develop motor dysfunctions.

ALS and FTD share several distinctive tracts at many levels.[209] Studies through CT and MRI demonstrated an involvement of brain beyond the motor regions, with progressive atrophy in frontotemporal lobes and limbic areas in ALS patients.[210] Frontotemporal hypoperfusion, as well as mild whole brain volume loss, have been also seen.[211,212] Among the several links that connect ALS with FTD, common pathological mechanisms are included. TDP-43- and FUS-positive inclusions have been detected in up to 90% of ALS and up to 50% of FTD patients, representing a common hallmark of both disorders.[151]

1.7.2 FTD Genetics

The pathogenic expansion in the *C9ORF72* gene represents the major genetic contributor in both diseases, accounting for up to 6% of sporadic and 25% of familial FTD cases.[213,214] To date, significant progresses have been made in unravelling the

genetics underneath the ALS/FTD *continuum*, showing a significant overlap (Figure 6). Indeed, besides some genes that are ALS-specific (i.e. *SOD1* and *VAPB*) or FTD-specific (i.e. *PGRN* and *MAPT*), many other genes, like *C9ORF72*, *TARDBP*, *SQSTM1*, *VCP*, *FUS*, *TBK1*, *CHCHD10* have been described to be causative of both diseases.[204]

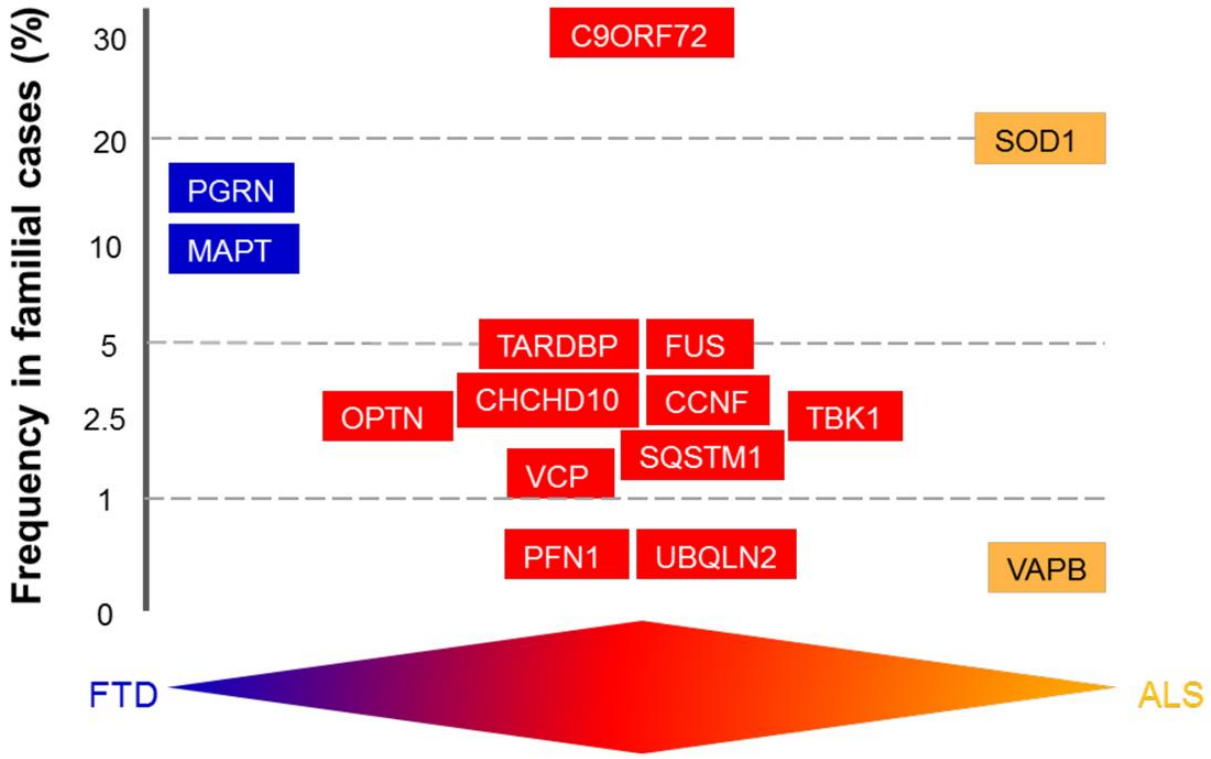

Figure 6| Schematic diagram illustrating the genetic overlap between ALS and FTD.[215]

1.8 Hereditary Spastic Paraplegia (HSP)

HSP is a syndromic designation for a clinically and genetically heterogeneous group of inherited neurodegenerative or neurodevelopmental disorders in which the main symptoms are lower limb spasticity and weakness. This lower limb impairment is caused by distal axonal degeneration that involves the longest axons of the corticospinal tracts due to spared molecular mechanisms.[216,217]

HSP can be classified, in various ways, according to the clinical phenotype, the pattern of inheritance, or pathophysiological molecular mechanism.[218] Based on the clinical phenotype, HSP can be classified in pure or complex forms. Pure HSP exhibits isolated pyramidal signs (spasticity, paraparesis or quadriparesis, brisk tendon reflexes, and extensor plantar responses) which can be associated with sphincter disturbances and deep sensory loss. In general, most HSP patients report a very slow progression that may start years before patient's complaint. In complex

HSP forms, spastic paraplegia is also associated with other neurological or non-neurological signs, including cerebellar dysfunction, cognitive impairment, psychiatric disturbances, axonal or demyelinating peripheral neuropathy, epilepsy, myopathic features, extrapyramidal features (Parkinsonism, dystonia). Non-neurological manifestations of complicated HSP are broad and heterogeneous, including dysmorphic features, ophthalmological abnormalities, and orthopedic abnormalities.[216]

1.8.1 HSP Genetics

HSP has a prevalence of 1.8 cases every 100.000 individuals[219] and it is characterized by an extremely diversified genetic contribution. Indeed, 72 different disease loci and 57 causative genes have been identified so far (Table 2). The modern genetic classification of the different forms of HSP is based on the chromosomal locus, causative mutation (if known) and pattern of inheritance. HSP can be inherited in different mendelian patterns, indeed either AD, AR, X-linked and maternal (mitochondrial) patterns of inheritance have been observed.[216,217] HSP age of onset ranges from early childhood to above 70 years of age and it depends mainly on the underlying genetics, however, variability between family members carrying the same mutation has been reported.

HSP	Gene (RefSeq NM)	Protein (RefSeq NP)	Disease (OMIM)	Inherit
SPG1	L1CAM (NM_001278116.2)	L1 cell adhesion molecule (NP_001265045.1)	Corpus callosum, partial agenesis of (MIM:304100)	XLR
			CRASH syndrome; MASA syndrome (MIM:303350)	XLR
			Hydrocephalus due to aqueductal stenosis (MIM:307000)	XLR
SPG2	PLP1 (NM_001128834.2)	Proteolipid protein 1 (NP_001122306.1)	Pelizaeus-Merzbacher disease (MIM:312080)	XLR
			Spastic paraplegia 2 (MIM:312920)	XLR
SPG3	ATL1 (NM_015915.5)	Atlastin GTPase 1 (NP_056999.2)	Neuropathy, hereditary sensory, type ID (MIM:613708)	AD
			Spastic paraplegia 3A (MIM:1826009)	AD
SPG4	SPAST (NM_014946.3)	Spastin (NP_055761.2)	Hereditary motor and sensory neuropathy, type 2C or Charcot-Marie-Tooth disease, type 2C (MIM:606071)	AD

SPG5	CYP7B1 (NM_004820.5)	Cytochrome P450, family 7, subfamily B, polypeptide 1 (NP_004811.1)	Bile acid synthesis defect, congenital, 3 (MIM:613812)	AR
			Spastic paraplegia 5A (MIM:270800)	AR
SPG6	NIPA1 (NM_144599.5)	NIPA magnesium transporter 1 (NP_653200.2)	Spastic paraplegia 6 (MIM:600363)	AD
SPG7	SPG7 (NM_003119.4)	SPG7 matrix AAA peptidase subunit, paraplegin (NP_003110.1)	Spastic paraplegia 7 (MIM:607259)	AD, AR
SPG8	KIAA0196/WASHC5 (NM_014846.4)	WASH complex subunit 5 (NP_055661.3)	Spastic paraplegia 8 (MIM:603563)	AD
			Ritscher-Schinzel syndrome 1 (MIM:220210)	AR
SPG9	ALDH18A1 (NM_002860.4)	aldehyde dehydrogenase 18 family member A1 (NP_002851.2)	Spastic paraplegia 9A (MIM:601162)	AD
			Spastic paraplegia 9B (MIM:616586)	AR
			Cutis laxa, Type 3 (MIM:616603; MIM:219150)	AD/AR
SPG10	KIF5A (NM_004984.4)	Kinesin family member 5a (NP_004975.2)	{Amyotrophic lateral sclerosis, susceptibility to, 25} (MIM:617921	AD
			Myoclonus, intractable, neonatal (MIM:617235)	AD
			Spastic paraplegia 10 (MIM:604187)	AD
SPG11	SPG11 (NM_025137.3)	SPG11 vesicle trafficking associated, Spatacsin (NP_079413.3)	Amyotrophic lateral sclerosis 5, juvenile (MIM:602099)	AR
			Charcot-Marie-Tooth disease, axonal type 2X (MIM:616668)	AR
			Spastic paraplegia 11 (MIM:604360)	AR
SPG12	RTN2 (NM_005619.5)	reticulon 2 (NP_005610.1)	Spastic paraplegia 12 (MIM:604805)	AD
SPG13	HSPD1 (NM_002156.5)	heat shock protein family D (Hsp60) member 1 (NP_002147.2)	Spastic paraplegia 13 (MIM:605280)	AD
			Leukodystrophy, hypomyelinating, 4 (MIM:612233)	AR
SPG14	Unknown			AR
SPG15	ZFYVE26 (NM_015346.4)	Zinc finger FYVE-type containing 26 (NP_056161.2)	Spastic paraplegia 15 (MIM:270700)	AR
SPG16	Unknown			XLR
SPG17	BSCL2 (NM_001122955.3)	BSCL2 lipid droplet biogenesis associated, Seipin (NP_001116427)	Lipodystrophy congenital generalized type 2 (MIM:269700)	AR
			Encephalopathy progressive, +/- lipodystrophy (MIM:615924)	AR
			Neuronopathy, distal hereditary motor, 5A (MIM:600794)	AD
			Spastic paraplegia 17 (MIM:270685)	AD
SPG18	ERLIN2 (NM_007175.8)	ER lipid raft associated 2 (NP_009106.1)	Spastic paraplegia 18 (MIM:611225)	AR
SPG19	Unknown			AD

SPG20	SPART (NM_015087.5)	Spartin (NP_055902.1)	Troyer syndrome (MIM:275900)	AR
SPG21	SPG21 (NM_016630.7)	SPG21 abhydrolase domain-containing, maspardin (NP_057714.1)	Mast syndrome (MIM:248900)	AR
SPG22	SLC16A2 (NM_006517.5)	solute carrier family 16 member 2 (NP_006508.2)	?Spastic paraplegia 22	XLR
			Allan-Herndon-Dudley syndrome (MIM:300523)	XLR
SPG23	Unknown			AR
SPG24	Unknown			AR
SPG25	Unknown			AR
SPG26	B4GALNT1 (NM_001478.5)	beta-1,4-N-acetyl-galactosaminyltransfe rase 1 (NP_001469.1)	Spastic paraplegia 26 (MIM:609195)	AR
SPG27	Unknown			AR
SPG28	DDHD1 (NM_030637.3)	DDHD domain containing 1 (NP_085140.2)	Spastic paraplegia 28 (MIM:609340)	AR
SPG29	Unknown			AD
SPG30	KIF1A (NM_001244008.1)	Kinesin family member 1A (NP_001230937.1)	Mental retardation, autosomal dominant 9 (MIM:614255)	AD
			Neuropathy, hereditary sensory, type IIC (MIM:614213)	AR
			Spastic paraplegia 30, autosomal recessive (MIM:610357)	AR
SPG31	REEP1 (NM_001164730.1)	Receptor expression-enhancing protein1 (NP_001158202)	Spastic paraplegia 31 (MIM:610250)	AD
SPG32	Unknown			AR
SPG33	ZFYVE27 (NM_001385871.1)	zinc finger FYVE-type containing 27 (NP_001372800.1)	Spastic paraplegia 33 (MIM:610244)	AD
SPG34	Unknown			XLR
SPG35	FA2H (NM_024306.5)	Fatty acid 2-hydroxylase (NP_077282.3)	Spastic paraplegia 35, autosomal recessive (MIM:612319)	AR
SPG36	Unknown			AD
SPG37	Unknown			AD
SPG38	Unknown			AD
SPG39	PNPLA6 (NM_001166111.1)	Patatin like phospholipase domain containing 6 (NP_001159583.1)	Boucher-Neuhauser syndrome (MIM:215470)	AR
			Spastic paraplegia 39 (MIM:612020)	AR
			Neuronopathy, distal hereditary motor, 5B (MIM:614751)	AR
			Laurence-Moon (MIM:245800); Oliver-McFarlane (MIM:275400)	AR
SPG40	Unknown			AD
SPG41	Unknown			AD

SPG42	SLC33A1 (NM_004733.4)	solute carrier family 33 member 1 (NP_004724.1)	Spastic paraplegia 42 (MIM:612539)	AD
SPG43	C19ORF12 (NM_001031726.3)	chromosome 19 open reading frame 12 (NP_001026896.2)	Congenital cataracts, hearing loss, and neurodegeneration (MIM:614482)	AR
			?Spastic paraplegia 43 (MIM:615043)	AR
SPG44	GJC2 (NM_020435.4)	gap junction protein gamma 2 (NP_065168.2)	Neurodegeneration with brain iron accumulation 4 (MIM:614298)	AD/AR
			Spastic paraplegia 44 (MIM:613206)	AR
SPG45 SPG65	NT5C2 (NM_001351172.2)	5'-nucleotidase, cytosolic II (NP_001338101.1)	Leukodystrophy, hypomyelinating, 2 (MIM:608804)	AR
			Lymphatic malformation 3 (MIM:613480)	AD
			Spastic paraplegia 45 (MIM:613162)	AR
SPG46	GBA2 (NM_020944.3)	glucosylceramidase beta 2 (NP_065995.1)	Spastic paraplegia 46 (MIM:614409)	AR
SPG47	AP4B1 (NM_006594.5)	adaptor related protein complex 4 subunit beta 1 (NP_006585.2)	Spastic paraplegia 47 (MIM:614066)	AR
SPG48	AP5Z1 (NM_014855.3)	Adaptor-related protein complex 5, zeta 1 subunit (NP_055670.1)	Spastic paraplegia 48 (MIM:613647)	AR
SPG49	TECPR2 (NM_014844.5)	tectonin beta-propeller repeat containing 2 (NP_055659.2)	Spastic paraplegia 49 (MIM:615031)	AR
SPG50	AP4M1 (NM_004722.4)	adaptor related protein complex 4 subunit mu 1 (NP_004713.2)	Spastic paraplegia 50 (MIM:612936)	AR
SPG51	AP4E1 (NM_007347.5)	adaptor related protein complex 4 subunit epsilon 1 (NP_031373.2)	Spastic paraplegia 51 (MIM:613744)	AR
SPG52	AP4S1 (NM_007077.5)	adaptor related protein complex 4 subunit sigma 1 (NP_009008.2)	Stuttering, familial persistent, 1 (MIM:184450)	AD
			Spastic paraplegia 52 (MIM:614067)	AR
SPG53	VPS37A (NM_152415.3)	VPS37A subunit of ESCRT-I (NP_689628.2)	Spastic paraplegia 53 (MIM:614898)	AR
SPG54	DDHD2 (NM_015214.3)	DDHD domain containing 2 (NP_056029.2)	Spastic paraplegia 54 (MIM:615033)	AR
SPG56	CYP2U1 (NM_183075.3)	cytochrome P450 family 2 subfamily U member 1 (NP_898898.1)	Spastic paraplegia 56 (MIM:615030)	AR

SPG55	C12orf65 (NM_152269.5)	chromosome 12 open reading frame 65 (NP_689482.1)	Spastic paraplegia 55 (MIM:615035)	AR
SPG57	TFG (NM_006070.6)	trafficking from ER to golgi regulator (NP_006061.2)	Combined oxidative phosphorylation deficiency 7 (MIM:613559)	AR
			?Spastic paraplegia 57 (MIM:615658)	AR
SPG58	KIF1C (NM_006612.6)	kinesin family member 1C (NP_006603.2)	?Spastic paraplegia 58	AR
SPG59	USP8 (NM_005154.5)	ubiquitin specific peptidase 8 (NP_005145.3)	Spastic ataxia 2 (MIM:611302)	AR
			?Spastic paraplegia 59	AR
SPG60	WDR48 (NM_020839.4)	WD repeat domain 48 (NP_065890.1)	Pituitary adenoma 4, ACTH-secreting, somatic (MIM:219090)	
			?Spastic paraplegia 60	AR
SPG61	ARL6IP1 (NM_015161.3)	ADP ribosylation factor like GTPase 6 interacting protein 1 (NP_055976.1)	?Spastic paraplegia 61 (MIM:615685)	AR
SPG62	ERLIN1 (NM_006459.4)	ER lipid raft associated 1 (NP_006450.2)	Spastic paraplegia 62 (MIM:615681)	AR
SPG63	AMPD2 (NM_004037.9)	adenosine monophosphate deaminase 2 (NP_004028.4)	?Spastic paraplegia 63 (MIM:615686)	AR
SPG64	ENTPD1 (NM_001776.6)	ectonucleoside triphosphate diphosphohydrolase 1 (NP_001767.3)	Pontocerebellar hypoplasia, type 9 (MIM:615809)	AR
			?Spastic paraplegia 64 (MIM:615683)	AR
SPG66	ARSI (NM_001012301.4)	arylsulfatase family member I (NP_001012301.1)	?Spastic paraplegia 66	AR
SPG67	PGAP1 (NM_024989.4)	post-GPI attachment to proteins inositol deacylase 1 (NP_079265.2)	?Spastic paraplegia 67	AR
SPG68	FLRT1 (NM_013280.5)	fibronectin leucine rich transmembrane protein 1 (NP_037412.2)	Mental retardation (MIM:615802)	AR
			?Spastic paraplegia 68	AR
SPG69	RAB3GAP2 (NM_012414.4)	RAB3 GTPase activating non-catalytic protein subunit 2 (NP_036546.2)	?Spastic paraplegia 69	AR
SPG70	MARS (NM_001257372.2)	Methionyl-tRNA synthetase (NP_001244301.2)	Martsolf syndrome (MIM:212720)	AR
			Warburg micro syndrome 2 (MIM:614225)	AR
			?Spastic paraplegia 70	AR

SPG71	ZFR (NM_016107.5)	zinc finger RNA binding protein (NP_057191.2)	Charcot-Marie-Tooth disease, axonal, type 2U (MIM:616280)	AD
			Interstitial lung and liver disease (MIM:615486)	AR
			?Spastic paraplegia 71	AR
SPG72	REEP2 (NM_001271803.2)	receptor accessory protein 2 (NP_001258732.1)	?Spastic paraplegia 72 (MIM:615625)	AD/AR

Table 2| Table of the main genes involved in HSP. For each gene, gene name and the Refseq transcript ID, protein name and the Refseq ID, associated diseases (OMIM) and pattern inheritance has been described.[216]

Recently, advances in molecular diagnosis have revealed a large clinical and genetic overlap either within different HSP subtypes[220], and between HSP and other NDs such as hereditary ataxia, ALS[221], primary lateral sclerosis (PLS)[222], and CMT.[223] In 2005, Meyer and colleagues reported one patient with a non-fatal course of juvenile ALS and a long-term progression, who presented a missense mutation in *SPG4.*[224] Brugman and collaborators, in the same year, described a further missense mutation of *SPG4* in a patient with a rapidly progressive spinal and bulbar upper motoneuron syndrome that progressed to ALS.[225] Beyond *SPG4*, pathogenic variants in *SPG11* have been associated with AR-juvenile ALS disease and CMT;[221] missense mutations within *KIF5A* (associated toHSP and CMT2) have been found in ALS patients.[142] The high degree of genetic heterogeneity, together with wide spectra for age at onset, symptoms, progression and their clinical overlap with other neurological diseases, make it difficult to establish genotype-phenotype correlations.[226]

An increasing number of genes, primarily associated with other NDs, such as leukodystrophies, are sometimes considered as differential diagnoses for HSP. Actually, eight leukodystrophies genes have been reported to mimic HSP in some cases: ATP binding cassette subfamily D member 1 (*ABCD1*), adenosine deaminase RNA specific (*ADAR*), eukaryotic translation initiation factor 2B subunit epsilon (*EIF2B5*), gap junction protein alpha 1 (*GJA1*), interferon-induced with helicase C domain 1 (*IFIH1*), ribonuclease H2 subunit B (*RNASEH2B*), SAM and HD domain-containing deoxynucleoside triphosphate triphosphohydrolase 1 (*SAMHD1*) and tubulin beta 4A class IVa (*TUBB4A*).[227] Moreover, HSP and leukodystrophies have at least three causal genes in common: proteolipid protein 1 (*PLP1*), heat shock protein family D (Hsp60) member 1 (*HSPD1*), gap junction protein gamma 2 (*GJC2*).[226] Genes primarily linked to PD have also been recently associated with HSP: the ATPase cation transporting 13A2 (*ATP13A2*) gene, as well as loss-of-function

mutations in ubiquitin C-terminal hydrolase L1 (*UCHL1*) gene, which has also been found in spastic ataxia patients.[226]

1.9 Parkinson's disease (PD)

PD was first described in 1817 by James Parkinson.[228] After AD, it is considered the most common ND, characterized by the degeneration of specific neuronal populations in the brain, particularly the dopaminergic neurons of the *substantia nigra pars compacta* (SNpc), which helps to control voluntary movement.[228] The loss of the dopaminergic neurons leads to the motor symptoms that characterize PD patients: resting tremor, bradykinesia, postural instability, and rigidity. However, many patients complain also other nonmotor symptoms, including fatigue, depression, anxiety, sleep disturbances, constipation, bladder and other autonomic disturbances (sexual, gastrointestinal).[229]

PD has a prevalence of approximately 1% at age 65, which rises to nearly 5% at 85 years of age.[230] The mean age of PD diagnosis is in the seventh decade of life, but an estimated 3% of cases occur in individuals younger than age 50.[231] PD is a progressive disease and may last 10–20 years after diagnosis. It is generally partially treatable (mainly with dopamine-replacement therapy) for a few years after diagnosis, but this is commonly followed by years of decline during which there is no effective therapy, ultimately leading to premature death.[228]

1.9.1 PD Genetics

Although 10% of PD patients exhibit a positive familial history, most PD cases are represented by sporadic forms. For most of the twentieth century, genetic predisposition was thought to play a negligible role in the disease, but in the past 15 years, the identification of distinct genetic loci responsible for both dominant and recessive inherited forms of PD provided vital clues in understanding its molecular pathogenesis. To date, more than 20 PD-related genetic loci (the PARK loci) and 18 pathogenic genes associated with these PARK loci have been identified in familial or sporadic PD (Table 3).

PD	Gene (RefSeq NM)	Protein (RefSeq NP)	Disease (OMIM)	Inherit
PARK1/4	SNCA (NM_000345.4)	Synuclein alpha (NP_000336.1)	Dementia, Lewy body (MIM:127750)	AD
			Parkinson disease 1 (MIM:168601)	AD
			Parkinson disease 4 (MIM:605543)	AD
PARK2	PRKN (NM_004562.3)	Parkin RBR E3 ubiquitin-protein ligase (NP_004553.2)	Parkinson disease, juvenile, type 2 (MIM:600116)	AR
PARK3		Unknown		AD
PARK5	UCHL1 (NM_004181.5)	Ubiquitin C-terminal hydrolase L11 (NP_004172.2)	{?Parkinson disease 5, susceptibility to} (MIM:613643)	AD
			Spastic paraplegia 79 (615491)	AR
PARK6	PINK1 (NM_032409.3)	PTEN induced putative kinase 1 (NP_115785.1)	Parkinson disease 6, early onset (MIM:605909)	AR
PARK7	PARK7 (NM_007262.5)	Parkinsonism associated deglycase (NP_009193.2)	Parkinson disease 7, early-onset (MIM:606324)	AR
PARK8	LRRK2 (NM_198578.4)	Leucine-rich repeat kinase 2 (NP_940980.4)	{Parkinson disease 8} (MIM:607060)	AD
PARK9	ATP13A2 (NM_022089.4)	ATPase cation transporting 13A2 (NP_071372.1)	Kufor-Rakeb syndrome (MIM:606693)	AR
			Spastic paraplegia 78 (MIM:617225)	AR
PARK10		Unknown		RF
PARK11	GIGYF2 (NM_015575.4)	GRB10 interacting GYF protein 2 (NP_056390.2)	{Parkinson disease 11} (MIM:607688)	
PARK12		Unknown		RF
PARK13		Unknown		RF
PARK14	PLA2G6 (NM_003560.4)	Phospholipase A2 group VI (NP_003551.2)	Infantile neuroaxonal dystrophy 1 (MIM:256600)	AR
			Neurodegeneration with brain iron accumulation 2B (MIM:610217)	AR
			Parkinson disease 14 (MIM:612953)	AR
PARK15	FBXO7 (NM_012179.4)	F-box protein 7 (NP_036311.3)	Parkinson disease 15 (MIM:260300)	AR
PARK16		Unknown		RF
PARK17	VPS35 (NM_018206.6)	Vacuolar protein sorting 35 homolog (S. cerevisiae) (NP_060676.2)	{Parkinson disease 17} (MIM:614203)	AD
PARK18	EIF4G1 (NM_182917.4)	Eukaryotic translation initiation factor 4 gamma, 1 (NP_886553.3)	{Parkinson disease 18} (MIM:614251)	AD
PARK19	DNAJC6 (NM_014787.4)	DnaJ heat shock protein family (Hsp40) member C6 (NP_055602.1)	Parkinson disease 19a, juvenile-onset (MIM:615528)	AR
			Parkinson disease 19b, early-onset (MIM:615528)	AR

PARK20	SYNJ1 (NM_203446.3)	Synaptojanin 1 (NP_982271.3)	Epileptic encephalopathy, early infantile, 53 (MIM:617389)	AR
			Parkinson disease 20, early-onset (MIM:615530)	AR
PARK21	DNAJC13 (NM_001329126.2)	DnaJ heat shock protein family (Hsp40) member C13 (NP_001316055.1)	Parkinson disease 21 (MIM:616361)	AD
PARK22	CHCHD2 (NM_001256864.2)	Coiled-coil-helix-coiled-coil-helix domain containing 2 (NP_001243793.1)	Parkinson disease 22 (MIM:616710)	AD
PARK23	VPS13C (NM_020821.3)	Vacuolar protein sorting 13 homolog C (S. cerevisiae) (NP_065872.1)	Parkinson disease 23, early onset (MIM:616840)	AR
PARK24	USP24 (NM_015306.3)	ubiquitin specific peptidase 24 (NP_056121.2)	{Parkinson disease 10}	
PARK suscep	GBA (NM_001005742.2)	Glucosylceramidase beta (NP_001005742.1)	{Lewy body dementia, susceptibility to} (MIM:127750)	AD
			{Parkinson disease, late-onset, susceptibility to} (MIM:168600)	AD
			Gaucher disease (MIM:608013; MIM:230800; MIM:230900; MIM:231000; MIM:231005)	AR

Table 3| Table of the main genes involved in PD. For each gene, gene name and the Refseq transcript ID, protein name and the Refseq ID, associated diseases (OMIM) and pattern inheritance has been described.

Clinical overlapping among ALS and PD has been firstly reported in 1973, in four patients in whom familial or sporadic parkinsonism anticipated ALS, without cognitive impairment or dysautonomia.[232] Subsequently, the association between PD and ALS has been proposed as a syndrome characterized by the concomitant presence of these two disorders without dementia or dysautonomia. The presence of bradykinesia associated with rigidity or tremor occurs with a frequency that ranges from 5 to 17% in ALS patients.[233,234] However, studies on the occurrence of PD features in ALS are lacking, and genetic correlates have not been fully explored.

1.10 Overlapping between neurodegenerative diseases

Despite differences in the prevalence of clinical manifestations, NDs share common features such as the increased incidence with age, chronic and progressive nature, selective brain mass loss, and damage of the network of synaptic connections.[235,236] Moreover, pathological and molecular mechanisms, including misfolding, oligomerization, and accumulation of proteins in the brain, mitochondrial dysfunction, oxidative stress, and DNA damage are shared among NDs.[236,237] Neurodegenerative

overlap may underlie true comorbidity of two or more fully developed NDs or a specific ND which is also characterized by neuropathological deposits belonging to another neuropathological disorder. Such comorbidities may end up in a more severe phenotype, with faster progression and atypical presentations.[238] Conventionally, neurological disorders are diagnosed based on clinical features such as symptom-onset and disease course, and on the characterization of physical signs to assess the affected neuronal population. However, the clinical diagnosis of NDs based on the phenotype remains challenging in heterogeneous conditions with overlapping symptoms; even in hereditary diseases with an explicit genetic contribution, due to a poor genotype-phenotype correlation. The genetic contribution in NDs like ALS may be crucial in the disease onset. Indeed, a multistep hypothesis has been suggested to explain the late and apparently sporadic onset of ALS.[239] This hypothesis relies on six 'steps' required to trigger ALS onset, where genetic predisposition, environmental exposures and other unknown molecular alterations are considered. Within this hypothesis, known genetic mutations in genes strictly associated to ALS but also referred to other NDs such as FTD, PD and HSP as well as to other neuromuscular diseases may account for multiple steps. Hence, genetic studies of both clearly pathogenic variant and genetic modifiers play a key role in the context of ALS.

2. Pompe disease

Pompe disease, also known as glycogenosis type 2 or acid maltase deficiency, is a rare autosomal recessive monogenic disorder caused by biallelic mutations in the acid α-glucosidase (*GAA*) gene, a lysosomal enzyme involved in glycogen break down to produce glucose, one of the main energetic sources for muscle contraction. GAA deficiency leads to lysosomal impaired glycogen breaks down, with consequent lack of glucose formation and glycogen accumulation in multiple tissues, with cardiac and skeletal muscles being the most severely affected. This ultimately results in myocyte disruption and muscle weakness.

The first description of this pathology was made in 1932 by Johannes Cassianus Pompe, diagnosing idiopathic hypertrophic cardiomyopathy in a girl of 7 month of age. Interestingly, the cardiac muscle of the girl presented an abnormal accumulation of glycogen.[240] However, the knowledge about this disorder remained vague until the discovery of lysosomes as cellular organelles in 1955[241] and the discovery made by Henri G. Hers in 1963, who linked the deficiency of the lysosomal GAA enzyme to the disease.[242]

The incidence of the disease is highly heterogeneous, varying from 1/14,000 (Afro-Americans) to 1/ 600,000 (Portugal) depending on geographic regions and the diagnostic criteria utilized.[243,244] In Europe, it has been estimated that the prevalence of PD is 1:283,000, however, it may be higher due to misdiagnosis or non-diagnosis.[245]

Traditionally, Pompe disease has been subdivided into two major phenotypes – infantile-onset Pompe disease (IOPD) and late-onset Pompe disease (LOPD). The two phenotypes depend on i) age of onset, ii) organ involvement, iii) severity, and iv) rate of progression. Usually, IOPD is more severe compared to LOPD. Indeed, the earlier the onset of manifestations, the faster the rate of progression. IOPD arises within the first year after birth and is characterized by hypotonia and generalized muscle weakness (96% of cases), feeding difficulties and failure to thrive (53% to 57%), cardiomegaly (92%, commonly hypertrophic cardiomyopathy), hepatomegaly (90%) and macroglossia (62%).[246] Without therapeutic interventions, death commonly occurs in the first two years of life from cardiopulmonary insufficiency.[247]

When the onset of symptoms occurs after 12 months of age, we referred to LOPD. LOPD is less severe and it is characterized by slow progression, the age of onset can span from one year of age to late adulthood and usually does not involve the

cardiovascular system. Instead, LOPD involves skeletal muscles and the respiratory system. These patients become wheelchair and/or ventilator-dependent, and respiratory insufficiency is the leading cause of death.[248] The explanation underlying the arise of the two different types of diseases seems to rely on the deficiency of the enzymatic activity displayed by the patient. Specifically, LOPD patients have < 30% of enzymatic activity, while IOPD less than 1%.[249] Enzymatic deficiency ultimately relies on the type of mutation that occurs within the *GAA* gene.

2.1 Genetic bases

The *GAA* gene is localized to chromosome 17q25.2-q25.3, spans approximately 28 kb and comprises 20 exons.[250] The first exon is non-coding and contains the 5' untranslated sequences. Additionally, it is separated from the second one (which is coding) by a large intron of approximately 2.7 kb. Usually, the first non-coding exons which are far from the coding ones, take part to the promoter region and are characteristic of housekeeping genes.[251] The gene encodes for a precursor protein of 952 amino acids which undergoes post-translational modifications and proteolytic cleavage during its transport to the lysosome, to finally end up in a 70-kDa mature form.[252]

To date, 582 mutations throughout the whole gene have been reported at HGMD-http://www.hgmd.cf.ac.uk/ac/. Missense mutations are the most frequent (51%) followed by small deletions (14,9%), splicing variants (12.7%), nonsense variants (8.8%), small insertions/duplications (6.0%), gross insertions/deletions (3.3%), small indels (2.2%) and complex rearrangements (0.9%) (Figure 7).[253] Only one variant in the regulatory region has been described so far.

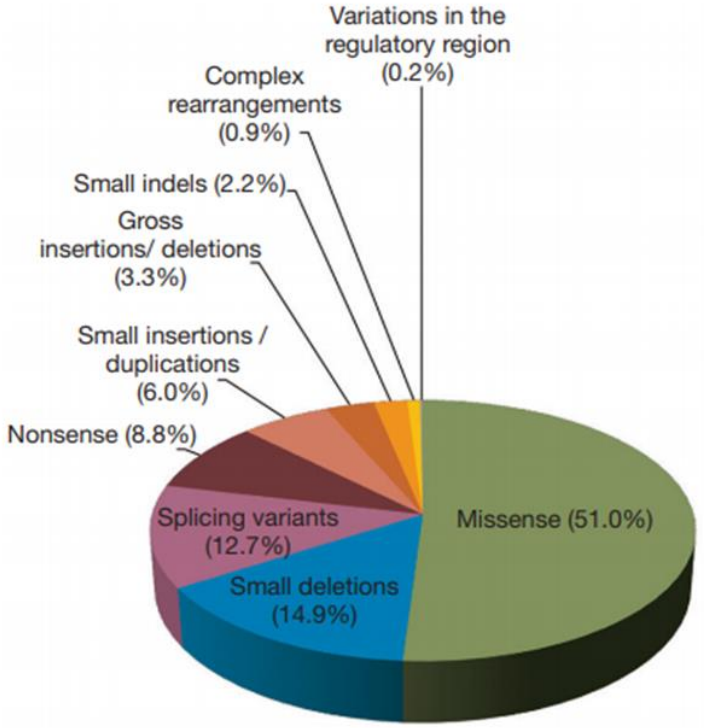

Figure 7| Frequency of the different types of mutations reported in the HGMD-http://www.hgmd.cf.ac.uk/ac/.[253]

Although *GAA* mutations are private or present in few families, the c.-32-13T>G splice mutation is the most common mutation in the Caucasian population, with a frequency that ranges from 40% to 70%, depending on the population studied. Despite the majority of mutations are IOPD- or LOPD-specific, the c.-32-13T>G has been detected in both forms.[254,255] This mutation affects the splicing efficiency, leading to the formation of a percentage of transcripts that lack exon 2, where the ATG codon is located, ending up in a non-functional protein. Thus, only a small portion of transcripts derived by the altered allele is finally translated into a functional protein. Therefore, the resulting variable levels of enzymatic activity may explain the delay of the onset of the disease in individuals who carry the c.-32-13T>G mutation.[256] As said, besides the c.-32-13T>G mutation, most mutations are private and only a few variants are recurrent. An example of other recurrent mutations is represented by the c.525delT and the c.2481+102_2646+31del variants, also known as “Dutch mutations”, that have been detected with a frequency of 35% and 31%, respectively in Dutch patients.[257] The latter causes the deletion of the entire exon 18 (Δ18) leading to a shift in the reading frame and to the insertion of an anticipated stop codon, eventually resulting in a truncated protein. The c.525delT mutation leads to a frameshift and to the insertion of a stop codon too. Moreover, nonsense-mediated

decay has been observed for this variant with no protein formation.[258] Due to the impact on the protein product, these mutations have been classified as severe and are associated with IOPD. Indeed their frequency in Dutch LOPD patients is only 15% and 8% respectively.[259]

2.1.1 Functional impact of GAA mutations

While it is quite easy to predict the impact of nonsense mutations or deletions/insertions that cause a shift in the reading frame on the function of the GAA enzyme, it is not that simple to understand the consequences of missense mutations or variants affecting sequences involved in the splicing process on the protein product. Although *in silico* prediction tools have been developed to predict the effect of such mutations on the transcript, these tools still cannot substitute functional studies, which remain the gold standard to determine their pathogenetic nature and to gain insights on their severity so far.[260] The severity of a particular mutation is defined based on the quantity and quality of processed and unprocessed GAA protein, and residual enzymatic activity, compared to the wildtype protein, in cells and media.[261] Kroos and colleagues proposed a classification of GAA mutations based on their severity and comprehends six classes: A) “very severe”; B) “potentially less severe”; C) “less severe”; D) “potentially mild”; E) “presumably nonpathogenic”; F) “nonpathogenic”.[261] In class A), no molecular forms of GAA can be detected by immunoblotting, and no residual enzymatic activity can be measured; in classes B), C) and D), the levels of processed and unprocessed forms of GAA are lower than normal ones, leading to a residual enzymatic activity ranging from 0 to 30%; in classes E) and F), quantitatively and qualitatively normal levels of GAA products are detected, but with an enzymatic activity higher than 30% and 60% respectively.[262] Class A) and B) are the most represented groups and comprise the 32,8% and 24,9% of mutations respectively, followed by class F) (16,7%), C) (5,4%), D) (3,9%), and E) (1,1%). Additionally, 15,2% of mutations have not been classified yet and are considered “unknown” (Figure 8).

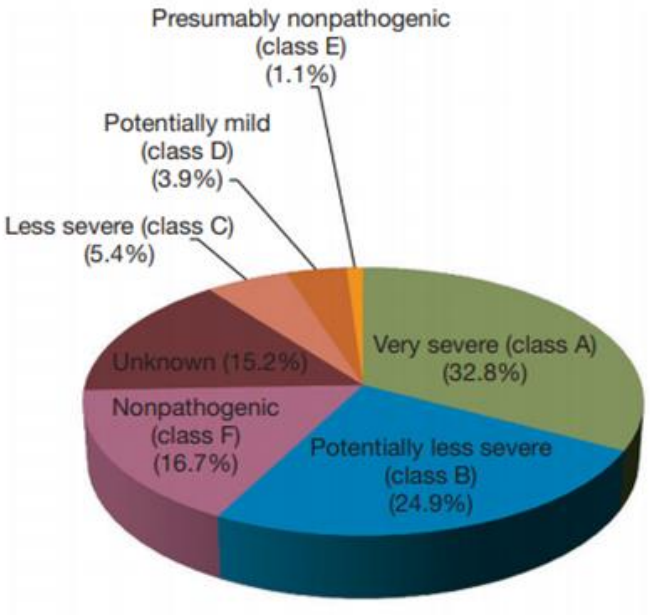

Figure 8| Schematic representation of the six classes of GAA mutations.[253]

2.1.2 Genotype-phenotype correlation

As for most genetic diseases, also in the case of Pompe disease there is not a strict correlation between the genotype of the patients and their phenotype, however, few assumptions can be done. As mentioned, patients affected by IOPD carry severe mutations that completely abolish the transcript expression or end up in non-functional proteins (Classes A and B). Conversely, LOPD patients harbour mild mutations at least in one allele, with the c.-32-13T>G mutation being the most common. Apart from these general considerations, most *GAA* mutations are private and are found in compound heterozygosity. To elucidate genotype-phenotype correlations, many authors have analysed the phenotype of patients carrying the c.-32-13T>G mutation in association with a "null" variant, thus that does not produce a functional protein with enzymatic activity. Also in this case, patients presented wide variability in terms of residual activity, age at onset, and disease progression.[263,264] Furthermore, variable clinical manifestations have been reported in siblings carrying this genotype.[265] All these reasons suggest that secondary factors, either genetic and non-genetic, may be responsible for the high clinical variability among patients. For instance, a polymorphism in the angiotensin I-converting enzyme (ACE) has been proposed as a possible genetic modifier, however, studies report controversial results.[266–268] Recently, the c.510C>T variant has been proposed as a genetic modifier of the age at onset, in compound heterozygous or homozygous for the c.-32-

13T>G variant. Indeed, if the two variants reside on the same allele, they modulate the splicing pattern of the mutated transcript reducing the amount of wildtype mRNA, eventually reducing the residual GAA activity.[269]

2.2 Pathogenesis

The GAA enzyme is involved in the breakdown of glycogen and a deficiency of this enzyme leads to glycogen accumulation within lysosomes of multiple tissues, especially in the cardiac tissue and in skeletal muscles. Glycogen-filled lysosomes eventually break disrupting cells, particularly myofibrils.[270,271] However, this pathological event does not take into consideration other secondary side effects derived by the accumulation of unmetabolized substrates in the lysosomes. Recently, some pathogenic mechanisms, such as autophagy, calcium homeostasis, oxidative stress, and mitochondrial abnormalities, have been described in Pompe disease.[272] For instance, micro-autophagy, which is the direct invagination of the lysosomal membrane to brings the cargo into the lysosomal lumen, has been seen to be profoundly dysregulated in Pompe disease,[273] being an important contributor to muscle weakness and it has been seen to be involved also in incomplete response to treatment (see section 2.3).[274] Moreover, since damaged mitochondria are cleared through the autophagic pathway, autophagy has been directly linked to mitochondrial abnormalities observed in muscle biopsies in a preponderant number of Pompe disease patients.[271,275]

2.3 Diagnosis and management

The increased awareness of the disease has allowed to improve the diagnosis, thus reducing the diagnostic delay and identifying the patients with mild signs and symptoms such as myalgia, fatigue or isolated hyperCKemia. Recently, the European Pompe Consortium (EPOC) listed some recommendations on the diagnosis and management of LOPD, defining the GAA enzymatic activity in dried blood spots (DBS) as a rapid and appropriate first-line diagnostic test.[276] Actually, it is considered the gold standard for the diagnosis. In LOPD, the residual enzyme activity does not exceed 30% of the normal activity.[277] Diagnosis may be further confirmed by assessing the GAA enzymatic activity in other tissue samples such as leucocytes, fibroblasts, or skeletal muscle and/or by performing the genetic screening of the *GAA* gene.[278] The assessment of vacuoles full of glycogen in lymphocytes has been also

suggested as a useful tool for Pompe diagnosis.[279] Eventually, muscular biopsy to evaluate myopathic changes with vacuolated myofibers and glycogen accumulation has been considered an important tool in LOPD diagnosis.[280] A summary of the current diagnostic methods is reported in the table below.

Method	Time	Invasive	False –	False +	Pros	Cons
Biochemistry						
Dried blood spot (DBS)	2–4 days	No	Yes	Yes	Cheap, fast, easy to perform, very useful as a screening tool	Biochemical and/or genetic confirmation
Fibroblasts	4–6 weeks	Yes (minimally)	No	No	Feasible, precise results	Long waiting period
Leucocytes	2–4 days	No	No	No	Easy to perform	Not always precise
Skeletal muscle	2–4 days	Yes	No	No	Feasible, precise results, moderate waiting period	Only in expert centers, usually in adults
Morphology						
Blood smear	Immediate	No	No	Yes	Cheap, fast, suggestive	Few studies, confirmation needed
Muscle morphology	2–4 weeks	Yes	Yes	Yes	If positive, strongly suggestive	Only in expert centers, confirmation needed
Genetic analysis						
GAA sequencing	Some weeks	No	No	No	Highly reliable if two pathogenic mutations	If not conclusive (only one mutation) biochemical confirmation needed

Table 4| Current diagnostic methods for LOPD diagnosis.[281]

Other investigations including laboratory examinations of different enzymes such as the creatine kinase (CK), aminotransferase (AST), alanine aminotransferase (ALT), and lactate dehydrogenase (LDH), which result often elevated in LOPD patients,[282,283] as well as pulmonary function evaluation by measuring maximum inspiratory pressure (MIP), maximum expiratory pressure (MEP), forced vital capacity (FVC), and vital capacity (VC) has been considered.[284,285] MRI can also help in evaluating the extent and localization of muscle changes in patients and can also help in identifying the site for a muscle biopsy. Talking about differential diagnosis, diseases that may resemble LOPD include limb-girdle muscle dystrophy, Duchenne muscular dystrophy and Becker muscular dystrophy, facioscapulohumeral muscular dystrophy, scapula-peroneal syndromes, rigid spine syndrome, myasthenia gravis, polymyositis, fibromyalgia, chronic fatigue syndrome, and glycogenosis types V and VI.[286]

The only pharmacological treatment for LOPD is the Enzymatic Replacement Therapy (ERT) using the *rhGAA* enzyme, approved in 2010.[287] The goals of *rhGAA* enzyme therapy depend on the stage of disease progression and include: to slow down, stabilize or reverse disease progression, increasing survival; to improve

mobility, preserve motor function, delaying the need for walking assistance; to improve or maintain respiratory function, thus preventing the need for respiratory assistance; to improve or preserve patient independence and quality of life.[288,289] Studies demonstrated that the earlier the beginning of ERT the earlier the getting of the clinical outcomes desired.[290] Adverse events in patients are considered rare and are usually infusion-related reactions. These reactions are usually mild to moderate and respond positively to a reduction in infusion rate. The most serious and rarely observed side effects of ERT with *rhGAA* are life-threatening anaphylactic reactions, serious allergic reactions and immune-mediated reactions.[291]

3. Genetic testing in rare diseases

The advent of NGS has revolutionized also the genetic field of neurological disorders, allowing the identification of thousands of genetic variants simultaneously.[292] Through NGS, the list of genes responsible for neurological dysfunction is rapidly expanding, helping to fill the gap of missing heritability in many NDs.[293] Moreover, through the identification of new and sometimes common genes in different disorders, it has become clear that common pathways exist in the pathogenesis of neurodegeneration. Additionally, the observation of pleiotropy has emerged, with mutations in the same gene giving rise to diverse phenotypes, which further increases the complexity of phenotype-genotype correlation, in which allelic heterogeneity and phenotypic heterogeneity coexist.[292] For these reasons, NDs like ALS are now considered polygenic diseases in which several genetic factors predispose to and are responsible for the onset and the development of the disorder. Polygenic diseases are caused by the combined influence of two or more genetic variants with small or moderate effects that ultimately result in the exhibition of clinically heterogeneous phenotypes.[294] Understanding the role of all the genetic contributors in influencing a phenotype remains quite challenging.[295,296] In light of this, the employment of the WES technology becomes useful in many ways: i) it gives the possibility to screen serveral genes simultaneously, both genes strictly correlated to the disease and genes associated with other disorders, ii) it allows the reanalysis of patients' genetic data in a subsequent time, useful especially when new genes associated with the diseases are discovere, iii) to look for genetic modifiers or to variants significantly associated with the disease.

On the other side, mendelian disorders, like LOPD, account for up to 80% of all rare diseases and are monogenic.[297] However, despite their manageable pattern of inheritance, the genetic basis of more than half of all identified Mendelian diseases has not been fully elucidated.[298,299] For instance, even if the causal disease gene of LOPD is known, the variable phenotype displayed by patients implies possible genetic modifiers, posing a challenge in the diagnostic and patient management. In this context, large-scale genome-wide sequencing studies may give additional information concerning penetrance and expressivity of rare diseases and may be useful to identify potential genetic modifiers. Hence, in the context of Mendelian disorders, the use of the WES technology may be emploied for the investigation of variants that are able to modify the clinical manifestation of the disease and that may have also a role in the different response to therapy displayed by patients rather than to look for causative mutations.

In conclusion, genetics plays a key role in the study of either complex rare diseases such as ALS, but also of Mendelian disorders like in the case of LOPD, and the employment of next generation techniques such as WES may turn out a useful and versatile tool that allows to obtain huge amount of data to performs several genetic analyses.

Aim of the study

Amyotrophic Lateral Sclerosis

Amyotrophic lateral sclerosis is a neuromuscular disorder that affect motor neurons leading to respiratory failure and death in 2 – 4 years after diagnosis. The 90% of ALS patients have no family history, while in the 10% of cases ALS segregates within families and have a genetic cause. However, an important genetic contribution has been observed also in sporadic patients, in which the 60% of genetic heritability has been observed. From here the new concept of ALS as an oligogenic and polygenic disease in which the sum several variants with small effects in different genes may lead to disease onset. To date, numerous studies have been carried out for the detection of ALS genetic determinants, however the genetic landscape continues to shift as the number of genes associated with the disease risk, pathogenesis, and cellular processes, keeps on growing. Moreover, ALS clinical manifestation is complicated by both genetic and phenotypic with other NDs, such as FTD, HSP and PD, but also with other neuromuscular diseases.[235,236] Hence, the overlap between ALS with other NDs gives rise to a more intriguing genetic interplay, which may eventually lead to a challenging differential diagnosis.

Starting from what mentioned above, the principal aims of this work were:

1) the study of the genetic architecture of sporadic ALS patients, through the employment of the NGS technology. Specifically, we recorded not only the pathogenic variants in the already known genes associated with ALS, but also rare variants that confer susceptibility to the disease or risk factor variants either in ALS genes but also in other diseases such as FTD, HSP, PD and other neurodegenerative and neuromuscular diseases;
2) to investigate i) if variants already known to be linked to ALS were overrepresented in our cohort of patients and ii) if some patients of our cohort presented oligogenic bases by harbouring more than one known ALS-linked variants.

This work is the results of our laboratory experience which started from the employment of customized gene panels to end up with the utilization of WES technology and virtual panels. Hence, our cohort of 248 patients comprises three subgroups of respectively 60, 107 and 81 patients. While the last group has been investigated through WES, the first two has been screened for 97 and 172 genes respectively.

Late-onset Pompe disease

LOPD is a rare monogenic autosomal recessive disorder, caused by mutations in the *GAA* gene. Defects in this gene, which encodes for the lysosomal alpha-glucosidase, lead to glycogen storage in lysosomes, especially in muscles. Traditional molecular genetic testing for LOPD is based on Sanger sequencing analysis of the *GAA* gene to identify point mutations and small deletion/insertion, while large deletions/duplications, not detectable through direct sequencing, are usually poorly investigated. Thus, the first aim of this study was to investigate, through Multiplexed ligation-dependent probe amplification (MLPA) analysis, a subset of patients diagnosed with LOPD in which only one point mutation was detected.

LOPD is characterized by highly variable clinical presentation, also between patients within the same family and thus with the same genotype. To date, no evidence for environmental factors able to modulate the clinical variability has been reported, thus genetic factors other than pathogenic *GAA* mutations might play a role in determining it. In the optic to look for genetic modifiers in the LOPD context, in the second part of this project we performed:

1. Studies on the impact of three single nucleotide polymorphisms (SPNs) within the GAA gene on protein stability;
2. WES analysis on 30 selected LOPD patients to look for variants associated with LOPD compared to the general population, in genes belonging to pathways relevant for the pathogenesis of LOPD. These variants may be responsible of clinical variability among patients.

Materials and Methods

4. Patients recruitment and clinical evaluation

4.1 ALS patients

248 patients affected by ALS have been recruited from different national centers (IRCCS Fondazione Mondino; ASST degli Spedali Civili di Brescia; Nuovo Ospedale Civile S. Agostino Estense di Modena; Ospedale S. Carlo Borromeo) after obtaining informed consent (Protocol RC20010Z). The institute's Ethics Committee (Ethics Committee Pavia) approved the study, as well as by the Ethics Committee of the Hospitals caring for the patients included in the study, in accordance with the Declaration of Helsinki. All patients underwent clinical and neurologic examinations and diagnosed with probable or definite ALS using the El Escorial criteria.[2]

4.2 LOPD patients

90 patients who received a diagnosis LOPD were recruited by several centers throughout Italy, participating in the study. All patients signed informed consent and the study was first approved in September 2016 by the Ethical Committee of the University of Pavia, and subsequently by the Ethics Committee of the Hospitals caring for the patients included in the study. Diagnosis of LOPD was done according to the diagnostic criteria of the American Association of Neuromuscular & Electrodiagnostic Medicine.[1]

5. Statistical analysis of clinical variables

Clinical records of ALS patients were examined for four phenotypic features: sex (male/female), age at disease onset (years), site of onset (bulbar or spinal), neuropsychological involvement (no, ALSbi, ALSci, ALS-FTD), phenotype[3] (classic, bulbar, flail arm, flail leg, pyramidal, respiratory, PLMN, PUMN), other extra-MND features (parkinsonian, cerebellar, psychiatric), rate of progression (fast, intermediate, slow) at last follow-up.

Statistical analyses were performed in a pairwise fashion between all clinical variables to identify significant associations. A $\chi 2$ analysis was performed between sex and i) site of onset, ii) neuropsychological involvement, iii) phenotype and iv) rate of progression; between age at onset and i) neuropsychological involvement and ii) phenotype; between site of onset and i) neuropsychological involvement, ii)

phenotype and iii) rate of progression; between neuropsychological involvement and i) phenotype and ii) rate of progression; between phenotype and rate of progression. Welch's t-tests were performed between age of onset and sex, while ANOVA was performed between age of onset and both site of onset and rate of progression.

6. Extraction of DNA and quality control

Peripheral blood has been collected for each patient and DNA has been extracted using the automated extractor Maxwell® 16 Instrument (Promega). Briefly, leukocytes have been collected using the TKM1 buffer (10mM Tris-HCl, pH 7.6; 10 mM KC1; 2 mM EDTA; 4 mM $MgCl_2$) and Nonidet P-40 substitute (Sigma-Aldrich) in 1:5 ratio from peripheral blood respectively. The mixture has been vigorously shaken to favor the cells lysis and leukocyte pellet has been subsequently recovered by centrifugation for 10 minutes at 2600 rotation per minute (rpm) at room temperature (RT) washing two times with 10 mL of TKM1. Eventually, the pellet has been loaded in a Maxwell® cartridge for automated extraction.

After DNA extraction, the concentration and the quality of the nucleic acid were assessed using two different methods, depending on the downstream application:

- *NanoDrop ND1000 UV-Vis Spectrophotometer* (Thermo Scientific). The instrument automatically measures sample absorbance at 260, 280, and 230 nm of wavelength according to Lambert-Beer law ($A = \varepsilon\lambda IM$),[3] determining the absorbance values, the concentrations, and the ratios at 260/280 nm and 260/230 nm. A 260/280 ratio of ~1.8 and a 260/230 ratio between 2 and 2.2 are generally accepted as "pure" for DNA (Figure 9).

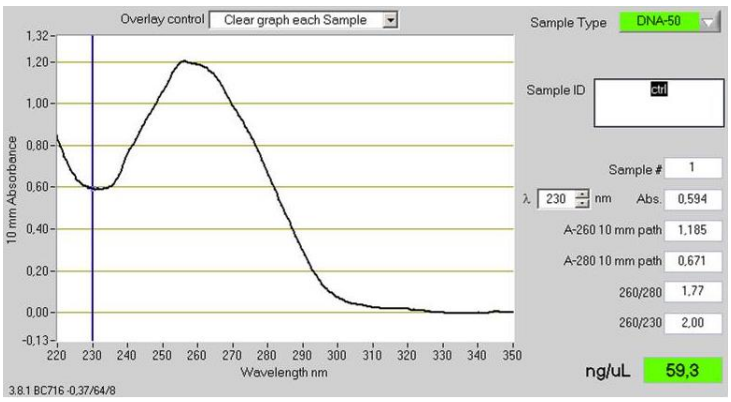

Figure 9| Example of sample quantification using NanoDrop UV-Vis Spectrophotometer. The shape of the curve relies on the qualitative and quantitative data showed on the right side of the graph.

- *Qubit fluorometric assay* (Invitrogen): Qubit is based on the intercalation of a fluorescent dye specific for double-strand DNA (dsDNA), thus avoiding the quantification of other contaminants such as degraded nucleic acid or free nucleotides. Qubit dsDNA Broad Range assay kit (2-1000 ng) and Qubit dsDNA High Sensitivity assay kit (0.2-100 ng) have been used in accordance with the different experiment settings (Figure 10).

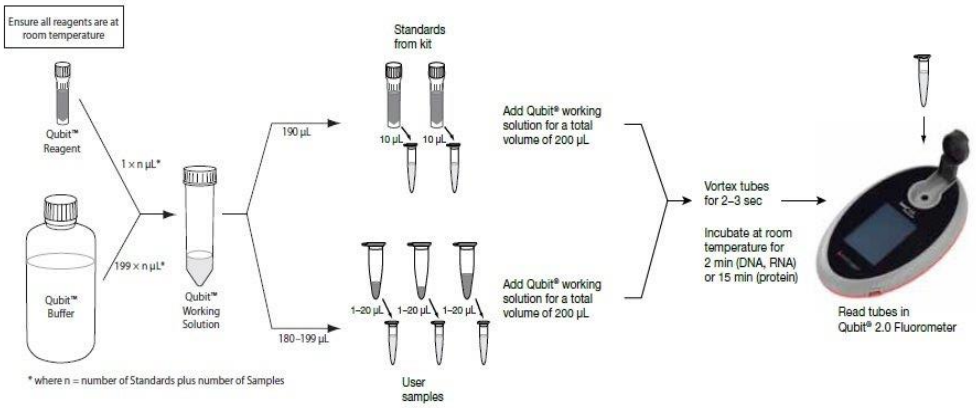

Figure 10| Qubit protocol for sample quantification.

7. Next Generation Sequencing (NGS)

NGS has been applied to screen a total of 248 ALS patients. Two different approaches have been employed for this purpose: i) Custom gene panels composed of 97 (P1) and 173 (P2) genes involved in the NDs allowed the screening of two different sub-cohorts of 60 and 107 patients respectively; ii) Whole Exome Sequencing (WES) instead, granted the chance to sequence all the coding DNA regions in another sub-cohort of 81 patients (P3; 277 genes). The different gene panel sizes reflect the advancement of the genetic screening for ALS patients adopted by our lab throughout the years (Figure 11).

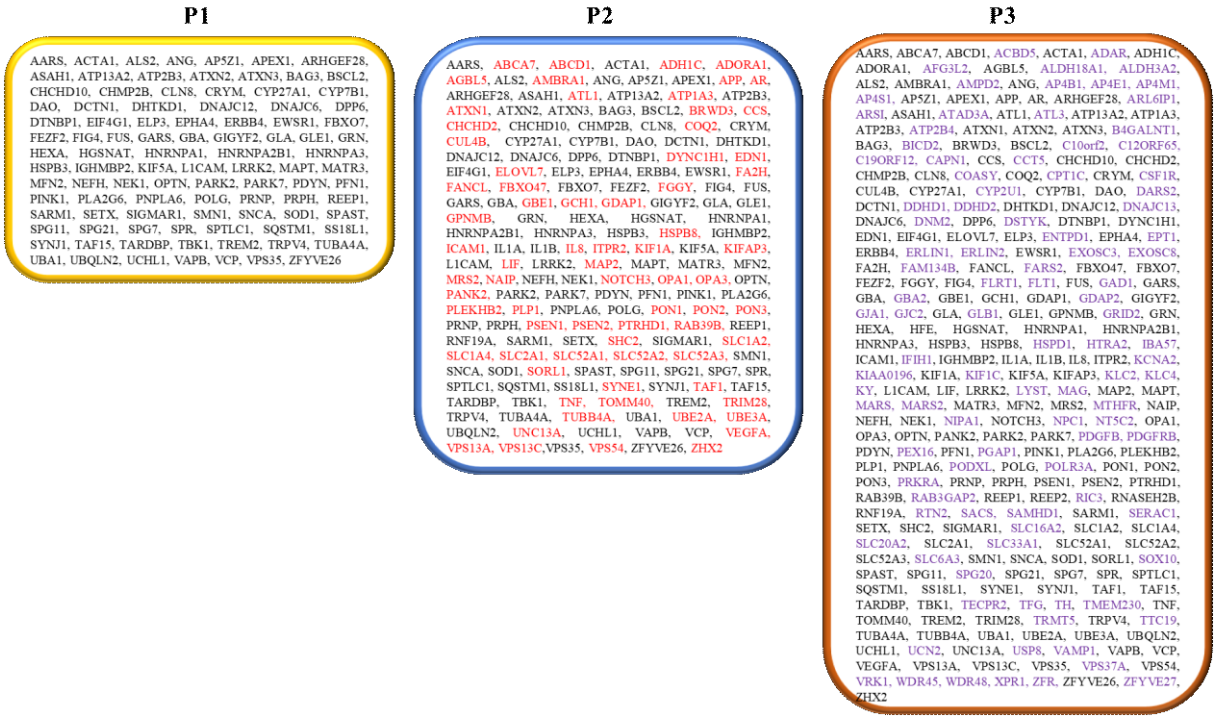

Figure 11| Lists of the genes screened by NGS. In red the genes added from Panel 1 to Panel 2, while in purple the genes added from Panel 2 to Panel 3.

Both approaches rely on the same DNA libraries preparation based on a Hybridization and Capture technology (SureSelect$^{XT/QXT}$, Agilent technologies; Figure 11), which requires a very low amount of starting DNA (50 ng).

A cohort of 30 LOPD were instead sequenced by using the SureSelect^{XT} Human All Exon V5 (Agilent Technologies) kit.

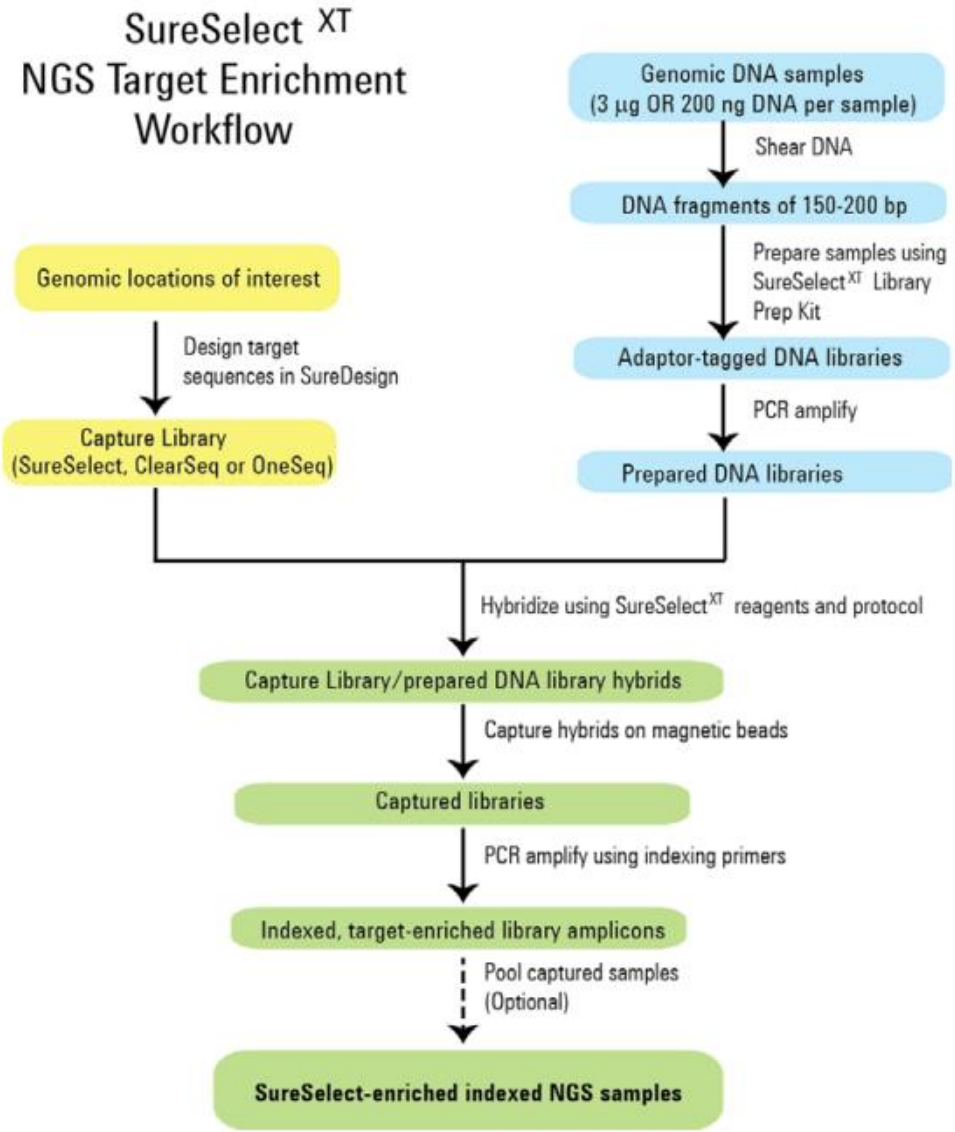

Figure 12| Target-enriched sequencing sample preparation workflow

7.1 Sample preparation

In this step, the gDNA has been enzymatically fragmented and adaptors have been added to the edges of the fragments in a single reaction. To do this, the following steps have been succeeded:

- Quantification of the gDNA as previously described and dilution to a final concentration of 25 ng/ µL;
- Enzymatic fragmentation with a pool of endonucleases to cut the DNA into 300 (base pair) bp fragments and ligation of the adaptor sequences to each generated fragment. Adaptors are instead necessary for the final sequencing run;

- Magnetic beads purification followed by amplification of the adapter-tagged libraries made by PCR reaction. PCR products are purified once again using magnetic beads;
- Evaluation of quality and quantity of the obtained DNA libraries, using the Tape Station D1000 kit (Agilent Technologies). It is important to verify that the electropherogram obtained shows the average size of the DNA fragments between 245 to 325 bp (Figure 12). A peak DNA fragment size significantly less than 245 bp may indicate too little gDNA in the fragmentation reaction and may be associated with increased duplicates in the sequencing data. Differently, a peak DNA fragment size greater than 325 bp may indicate too much gDNA in the fragmentation reaction and may be associated with decreased percent-on-target performance in sequencing results.

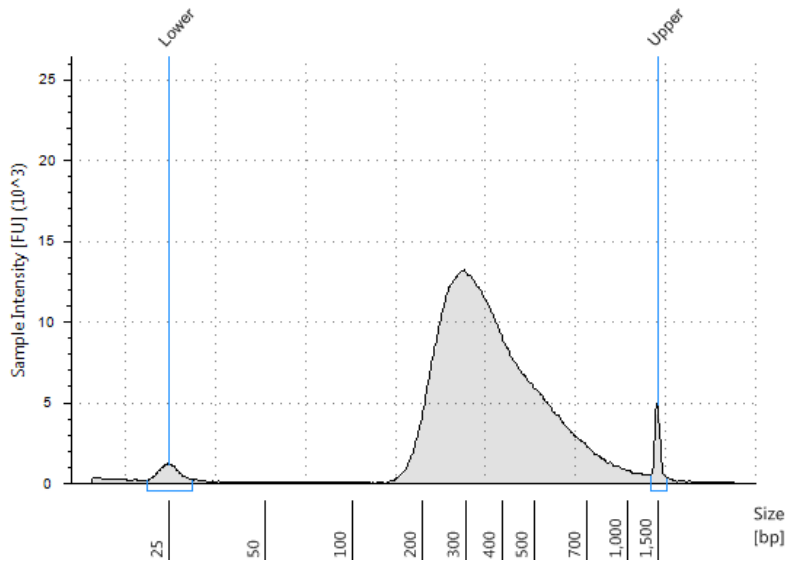

Figure 13| Representative sample electropherogram showing pre-capture analysis of amplified library DNA.

7.2 Hybridization and Capture

In this section, the prepared gDNA libraries have been hybridized with a target-specific Capture Library. According to the protocol, the input of the gDNA library used in this section varies according to the size of the Capture Library used. In our study, four Capture Libraries have been used: three “Neurodegeneration” Capture Libraries (< 3.0 Megabases) and the “Clinical Research Exome V2” (Exome) Capture Library (≥ 3.0 Mb) were used. Thus, the steps of this section of the protocol have been modified according to the different sizes of the Capture Libraries used.

The main steps of this section are the following:

- Normalization of the adaptor-tagged DNA libraries. For the "Neurodegeneration" Capture Libraries, libraries have been normalized to a final amount of 750 ng in 12 µL of the final volume. For the "Exome" Capture library, the input amount was instead 1500 ng in a final volume of 12 µL.
- Hybridization to the libraries of the designed Capture library containing the biotin-labeled probemix specific for the genetic regions to be analysed.
- Capture of the targeted molecules using the Dynabeads MyOne Streptavidin T1 (Thermo Fisher Scientific), which can recognize the biotinylated probemix used during hybridization. In this step, the targeted regions are retained and the fragments not of interest are discarded.

7.3 Indexing and Sample Processing for Multiplexed Sequencing

In this step, SureSelect-enriched DNA libraries have been PCR-amplified using the appropriate pair of dual, P7 and P5, indexing primers. Below the principal steps:

- Amplification of the SureSelect-enriched DNA libraries by PCR. In this step, a unique combination of two 8-bp indexes has been also added to each sample. These indexes will allow recognizing fragments coming from specific samples even when they will be pool together with fragments of other samples. PCR products are purified using magnetic beads.
- Qualitative and quantitative evaluation of DNA library made by Tape Station using High Sensitivity D1000 kit (Agilent Technologies). It is important to verify that the electropherogram shows the peak of DNA fragment size positioned between 325 and 450 bp. At this time, to be more precise, libraries have been quantified also using Qubit.
- Conversion of the quantification in ng/µL obtained with the Qubit into nM, using the Average Library Size obtained with the Tape Station and the following formula:

$$[nM] = \frac{\left[\frac{ng}{\mu L}\right]}{(660 * Caputered\ library\ average\ size)} * 10\text{^}6$$

- The pooling of the samples in equal amounts to a final concentration of 4 nM. To combine the correct amount of indexed-libraries, the following formula has been used:

$$The\ volume\ of\ Index = \frac{4\ nM * Final\ Volume\ of\ the\ Pool}{N°\ of\ indexes * Concentration\ of\ the\ sample}$$

Up to 24 samples hybridized with the Neurodegeneration Capture Libraries can be pooled together and sequenced on the Illumina MiSeq sequencer, while up to 6 samples hybridized with the Exome Capture Library can be loaded into the Illumina NextSeq500 sequencer per run.

7.4 Denature and Dilute Libraries

Preparation of the pool has been done following the “Denature and Dilute Libraries Guide" (Illumina) of the MiSeq for pools prepared with the Neurodegeneration Capture Library and the one of the NextSeq for pools prepared with the Exome Capture Library. Briefly, pools have been denatured using NaOH 0.2 N (Sigma-Aldrich) and subsequently diluted to the suited concentration, using the Hybridization Buffer (HT1) provided by Illumina. Neurodegeneration-derived pools have been diluted to a final concentration of 8 pM, while Exome-derived pools have been diluted to a final concentration of 1.2 pM. Once ready, diluted pools have been loaded into a cartridge and placed into the designed sequencer. The sequence reaction takes place on a flow cell, which is an optically transparent surface covered by short oligonucleotides complementary to the previously added adapter sequences. Adapted single-stranded DNAs are hybridized to the flow cell and amplified through the “bridge PCR” methodology. This process results in the formation of small clusters of identical DNA templates which can be sequenced all together using sequencing by synthesis application.
LOPD patients were instead sequenced on the HiSeq1000 platform (Illumina).

8. Bioinformatic Analysis

Data collected from NGS experiments for ALS patients have been analysed in order to identify single nucleotide substitutions and/or small insertions and deletions. The first step of the bioinformatic analysis has been carried out with the MiSeq software (Real-Time Analysis RTA v.1.18.54 and Casava v.1.8.2, Illumina, Inc., San Diego,

CA) by bioinformaticians. It performs the trimming of the adapters and the demultiplexing, such as the division of the reads obtained based on the dual indexes and thus based on the belonging sample, obtaining the FASTQ files. FASTQ files were then aligned, based on the Burrows-Wheeler transform, to the human reference genome (GRCh37) using BWA v7.5a, while BAM (*binary alignment map*) files have been obtained through the use of samtools v1.19 and v1.95 Picard-tool. GATK v3.1 has been used instead for the realignment of insertions/deletions (with RealignTargetCreator, IndelRealigner, and BaseRecalibrator) and for the variants calling (with UnifiedGenotyper) according to the Best Practices recommendations GATK. The alignment based on the Burrows-Wheeler transform has been performed with the software BWA v7.5a, while the BAM files have been obtained through the use of samtools v1.19 and v1.95 Picard-tool. GATK v3.1 has been used instead for the realignment of insertions/deletions (with RealignTargetCreator, IndelRealigner, and BaseRecalibrator) and for the variants calling (with UnifiedGenotyper) according to the Best Practices recommendations GATK, obtaining VCF (*Variant Call File*) files. VCFs have been eventually used as input files on the eVAI v0.7 (EnGenome) webtool for variant annotation and prioritization.

Both VCFs coming from the Neurodegeneration-targeted panels and Exome sequencing undergo discreet filtering based on:

- Quality filter: Variants must exceed the quality threshold established by software's algorithms, in order to discriminate potential false positive;
- The coverage (read depth) of each variant must be equal or higher than 15 reads; differently, all those variants identified by a coverage lower than 15 reads have been discarded;
- Minor allele frequency (MAF) must be lower than 1% in the 4 principal population databases (GnomAD, ExAC, 1000 GP, and ESP);
- Variants must fall within coding sequences, affecting splicing sites, and must be involved in alteration of the 3'-untranslated (UTR) or 5'-UTR. Intron variants, as well as intragenic and non-coding exon variants, were excluded.

Variants retained were further investigated using different software and databases:

- VarSome software (https://varsome.com/), to classify variants according to the guidelines of the American College of Medical Genetics (ACMG criteria) [4] (Table 5). According to the ACMG criteria, each variant is linked to a pathogenicity or benignity evidence parameter. These parameters are connected to different

variant characteristics and each one contributes in a different way to the calculation of the variants score;

- dbSNP and ClinVar, if variants have been already described and associated with specific diseases;
- GnomAD (https://GnomAD.broadinstitute.org/) database to observe the frequency of the variants in different populations and to know how the gene of interest is tolerant to synonymous, missense, and loss of function variants. Using GnomAD is also possible to verify if the variants found are present in healthy individuals and/or individuals which do not present a neurological phenotype.
- *In silico* prediction software (SIFT, Polyphen-2, Mutation Taster) to evaluate the physical-chemical consequence of an amino acid change on protein structure and functionality;
- Clustal W2 (https://www.ebi.ac.uk/Tools/msa/clustalw2/) to study the amino acid conservation among species (variants which affect an amino acid in a highly conserved position likely have a deleterious effect)
- Human Splicing Finder v3.1 (http://umd.be/Redirect.html) to study potential alteration of splicing sites;
- GTEx (https://gtexportal.org/home/index.html) and the human protein atlas (https://www.proteinatlas.org/) to verify the gene expression in different tissues.
- Mouse Genome informatics (MGI, http://www.informatics.jax.org/) to evaluate the possible effects due to lack of protein function in animal models.

Concerning the WES data of LOPD patients, reads were quality-filtered and aligned to the reference human genome sequence (GRCh37/hg19) with ISAAC aligner.[300] Variant calling was performed with ISAAC Aligner. Genomic variant annotation was carried out with VarSeq v1.4.5 (Golden Helix, Inc.) and only variants with a minimum quality score of 20 and a minimum read depth of 10X were included in the downstream analysis. Thereafter, we excluded variants reported in the ExAC and GnomAD databases with a population frequency above 5%. Finally, the manual inspection of the Bam files, by using Integrative Genomics Viewer (IGV), allowed us to evaluate the coverage and quality of the aligned reads.

	ACGM criteria	Meaning
Pathogenic	PVS1	Null variant (nonsense, frameshift, canonical ±1 or 2 splice sites, initiation codon, single or multiexon deletion) in a gene where LOF is a known mechanism of disease.
	PS1	Same amino acid change as a previously established pathogenic variant regardless of nucleotide change.
	PS2	De novo (both maternity and paternity confirmed) in a patient with the disease and no family history.
	PS3	Well-established *in vitro* or *in vivo* functional studies supportive of a damaging effect on the gene or gene product.
	PS4	The prevalence of the variant in affected individuals is significantly increased compared with the prevalence in controls.
	PM1	Located in a mutational hot spot and/or critical and well-established functional domain (e.g., the active site of an enzyme) without benign variation.
	PM2	Absent from controls (or at extremely low frequency if recessive) in Exome Sequencing Project, 1000 Genomes Project, or Exome Aggregation Consortium.
	PM3	For recessive disorders, detected in trans with a pathogenic variant
	PM4	Protein length changes as a result of in-frame deletions/insertions in a non-repeat region or stop-loss variants.
	PM5	Novel missense change at an amino acid residue where a different missense change determined to be pathogenic has been seen before.
	PM6	Assumed de novo, but without confirmation of paternity and maternity.
	PP1	Co-segregation with the disease in multiple affected family members in a gene definitively known to cause the disease.
	PP2	Missense variant in a gene that has a low rate of benign missense variation and in which missense variants are a common mechanism of disease.
	PP3	Multiple lines of computational evidence support a deleterious effect on the gene or gene product (conservation, evolutionary, splicing impact, etc.)
	PP4	The patient's phenotype or family history is highly specific for a disease with a single genetic aetiology.
	PP5	Reputable source recently reports variant as pathogenic, but the evidence is not available to the laboratory to perform an independent evaluation.
	BA1	Allele frequency is >5% in Exome Sequencing Project, 1000 Genomes Project, or Exome Aggregation Consortium.
	BS1	Allele frequency is greater than expected for the disorder.
	BS2	Observed in a healthy adult individual for a recessive (homozygous), dominant (heterozygous), or X-linked (hemizygous) disorder, with full penetrance expected at an early age.

	BS3	Well-established in vitro or in vivo functional studies show no damaging effect on protein function or splicing.
	BS4	Lack of segregation in affected members of a family.
	BP1	Missense variant in a gene for which primarily truncating variants are known to cause disease.
	BP2	Observed in trans with a pathogenic variant for a fully penetrant dominant gene/disorder or observed in cis with a pathogenic variant in any inheritance pattern.
	BP3	In-frame deletions/insertions in a repetitive region without a known function.
	BP4	Multiple lines of computational evidence suggest no impact on gene or gene product (conservation, evolutionary, splicing impact, etc.)
	BP5	Variant found in a case with an alternate molecular basis for disease.
	BP6	Reputable source recently reports variant as benign, but the evidence is not available to the laboratory to perform an independent evaluation.
	BP7	A synonymous (silent) variant for which splicing prediction algorithms predict no impact to the splice consensus sequence nor the creation of a new splice site AND the nucleotide is not highly conserved.

Table 5| Summary of the adapted ACMG pathogenic and benign criteria

For each gene analysed, pathogenic, likely pathogenic, and VUS (with uncertain significance) variants were retained and, validated in Sanger sequencing.

9. List of variants implicated in ALS

A list of 726 genetic variants reported in the literature as implicated in ALS (including ALS-linked mutations, functional risk alleles and disease-associated variants) within 25 ALS causative genes has been made based on the list already published by McCann and colleagues.[301] If a gene was missing in the list, a manual evaluation of all the resulting publications querying “gene name” and ‘amyotrophic lateral sclerosis’ on PubMed (https://www.ncbi.nlm. nih.gov/pubmed/) was done to look for variants. Variants were included in the list if they have an effect on the amino acidic sequence, such as missense, splicing variants or stop-gain, and if they were cited in the main text of the publication. Chromosome location (hg19), transcript accession number and cDNA and protein changes were recorded for each variant.

10. Identifying ALS-implicated and LOPD-implicated variants in our cohorts

ALS cohort: eVai software was interrogated to identify patients carrying the previously sorted ALS-implicated variants. For each variant found, frequency in our cohort has been recorded and compared to the MAF value from the ethnically matched non-neuro (ie, excluding individuals with neurological disorders) non-Finnish European (NFE) subset of the Genome Aggregation Database (GnomAD) by using contingency test. Among the 726 genetic variants also the HRE in the *C9ORF72* is present. The HRE was determined through microsatellite analysis (see section XXX). To determine whether any of the reported ALS-implicated variants were associated with ALS in our cohort, allele counts were compared between our cases (n = 248) and control individuals from the non-neuro NFE subset of GnomAD using Fisher's exact test (GraphPad Prism v7). Control genotyping data for the HRE in the *C9ORF72* gene was not available, therefore association testing was not completed for these variants.

LOPD cohort: For each variant found, frequency in our cohort has been recorded and compared to the MAF value from the ethnically matched NFE subset of the GnomAD database by using contingency test.

11. Association analysis

To determine whether any of the reported ALS-implicated variants were associated with our cohort of ALS patients, allele counts were compared between ALS cases (n = 248) and control individuals from the non-neuro NFE subset of GnomAD using Fisher's exact test. The same statistical analysis was performed also to study the association among the variants found through WES in the cohort of LOPD patients with the disease, comparing the allele count between LOPD patients (n = 30) and control individuals from the NFE subset of GnomAD population.

12. Polymerase Chain Reaction (PCR), detection and purification

The PCR reaction has been carried out by using the kit MyTaq® DNA Polymerase (Bioline), in a final volume of 25 µL, according to the following proportions:

Reagent	Volume
5x MyTaq Reaction Buffer	5 µL
Primer forward (10 µM)	1 µL
Primer reverse (10 µM)	1 µL
MyTaq® DNA Polymerase	0,25 µL
*gDNA or cDNA**	Variable
H2O	up to 25µL

*50-100 ng of gDNA or 10-20 ng of cDNA have been used.

PCR cycling conditions:

	Number of cycles	Temperature	Time
Initial denaturation	1	95 °C	4 minutes
Denaturation	35	95 °C	30 seconds
*Annealing**		T_m	30 seconds
Elongation		72 °C	30 seconds
Final Elongation	1	72 °C	7 minutes
Optional Step	1	4 °C	hold

*The optimal annealing temperature is dependent upon the primer sequences and is usually 2-5 °C below the lower Tm of the pair.

Per each PCR experiment, a negative control lacking the DNA template has been run in order to exclude eventual reagents' contaminations.

PCR products, lack of non-specific amplicons, and the negative control have been loaded on 1% Agarose together with the 100 bp DNA Ladder (Invitrogen) to verify the amplicon size and the electrophoretic run has been performed for 20 minutes at 130V. Eventually, bands have been detected using the Azure Biosystem c600 (Aurogene). Subsequently, amplicons have been purified (1-1.5 µL based on the intensity of detected band) using the illustra™ ExoProStar™ (GE Healthcare) which combines the activity of two enzymes, the Exonuclease I (Exo) and the Shrimp Alkaline Phosphatase (SAP), to remove primer dimers and unused dNTPs. PCR cycling condition as follow:

	Number of cycles	Temperature	Time
Enzymatic purification	1	37 °C	15 Minutes
Enzymes inactivation	1	80 °C	15 Minutes
Optional Step	1	4 °C	hold

13. Sanger sequencing

Through the automated Sanger sequencing, the DNA can be sequenced by generating fragments of different length through the random incorporation of 3'-dideoxynucleotide triphosphates (ddNTPs). ddNTPs lack the 3′ hydroxy- group, fundamental for the formation of the phosphodiester bond with the subsequent nucleotide. Thus, the incorporation of a ddNTP leads to the stop of the sequencing reaction. Moreover, each ddNTP is labeled with a different fluorescent tag that allow the identification (respectively green for adenine, blue for cytosine, black for guanine, and red for thymine).

The DNA polymerase can insert either normal deoxynucleotide triphosphates (dNTPs) or ddNTPs randomly. The ddNTPs incorporation blocks the chain elongation reaction and results in the production of fragments of variable lengths that can be resolved through capillary electrophoresis (CE).

The sequencing reaction has been performed in both directions, sense and antisense, using the BrilliantDye™ Terminator (v3.1) Cycle Sequencing kit (NimaGen).

The reaction mix has been set up as follow:

Reagent	Volume
BigDyeTM Terminator v3.1 Ready Reaction Mix	0,5µL
BigDyeTM Terminator v3.1 5X Sequencing Buffer	2µL
Dimethyl sulfoxide (DMSO)	0,5 µL
Primer forward or reverse (10 µM)	1µL
Purified PRC product	1 – 1,5 µL
H2O	Up to 10 µL

The sequencing thermal cycler program used is described below:

	Number of cycles	Temperature	Time
Initial denaturation	1	96 °C	1 minute
Denaturation	25	95 °C	10 seconds
Annealing		50 °C	5 seconds
Elongation		60 °C	1 minute
Optional Step	1	4 °C	hold

The obtained sequences have been purified with two different methods, based on the total number of ready reactions:

- DyeEx 2.0 Spin kit (Qiagen): Following the manufacturer's instructions, 5 µL of purified sequence has been used and 15 µL of formamide have been added for further run into the CE;
- Ethanol purification: 50 µL of Ethanol 100% and 2 µL of Sodium Acetate have been added to each sequence reaction for 20 minutes at -20 °C to allow the precipitation of nucleic acids. After precipitation, nucleic acids have been separated from the rest of the solution by centrifugation at 1900 x g for 30 minutes. The solution has been discarded with a sharp flicking motion and the precipitates have been centrifuged upside down for 1 minute at 700 x g to remove the

remaining drops. Subsequently, 150 µL of Ethanol 70% has been added to wash the nucleic acids and centrifuged at 1900 x g 15 minutes. The solution has been discarded with a sharp flicking motion and the precipitates have been centrifuged upside down for 1 minute at 700 x g. At the end, 20 µL of formamide has been added for further run into the CE.

Before loading the sequences into the ABI 3500 Genetic Analyser (Applied Biosystem) they have been undergone a step of denaturation (94 °C for 3 minutes) immediately followed by 2 minutes at -20 °C to stop and stabilize the sequences in single strand. Sanger sequencing relies on the CE methodology, which allows the migration of the sequence fragments based on their size. During the migration, the fluorescence of the fragments is read by a laser, which leads to an electromagnetic emission, captured by a detector that recognizes their wavelength. The result of all the captured emission is finally graphically represented by an electropherogram, where each peak identifies a different nucleotide in the sequence. The obtained sequences have been analysed using Sequencher 4.8 (Genecodes).

14. Microsatellite analysis

Since the detection of repeat expansions is not possible by NGS, microsatellite analysis has been performed to study the HRE in the *C9ORF72* gene. A repeat-primed PCR, according to the DeJesus-Hernandez et al. protocol that consists of two primers external to the HRE (one labeled with a fluorophore) and one internal primer homologous to the $(4G2C)_n$ sequence has been done.[5] The PCR reaction has been carried out by using the FastStart PCR Master and FastStart Taq Polymerase,dNTPack (Roche) in a final volume of 28 µL, according to the following proportions:

Reagent	Time
FastStart PCR Master	14 µL
*Primer Mix**	2 µL
DMSO	2 µL
1M betaine	5 µL
7-deaza-2-deoxy GTP (5 mN)	1 µL
MgCl2 (25 mM)	1 µL
DNA (20 ng/ µL)	3 µL
Total Volume	28 µL

*Primer Mix	Primer Sequence	Volume
F1 (100 µM)	**6-FAM-**ATGCGCTAGAGGCGAAAGC	20 µL (20 µM)
R (100 µM)	TACGCATCCCAGTTTGAGACGGGGGCC GGGGCCGGGGCCGGGG	10 µL (10 µM)
Tail (100 µM)	TACGCATCCCAGTTTGAGACG	20 µL (20 µM)
H2O		50 µL

The sequencing thermal cycler program used is described below:

	Number of cycles	Temperature	Time
Initial denaturation	1	95 °C	15 Minutes
Denaturation	2	94 °C	1 Minute
Annealing		70 °C	1 Minute
Elongation		72 °C	3 Minutes
Denaturation	3	94 °C	1 Minute
Annealing		68 °C	1 Minute
Elongation		72 °C	3 Minutes
Denaturation	4	94 °C	1 Minute
Annealing		66 °C	1 Minute
Elongation		72 °C	3 Minutes
Denaturation	5	94 °C	1 Minute
Annealing		64 °C	1 Minute
Elongation		72 °C	3 Minutes
Denaturation	6	94 °C	1 Minute
Annealing		62 °C	1 Minute
Elongation		72 °C	3 Minutes
Denaturation	7	94 °C	1 Minute
Annealing		60 °C	1 Minute
Elongation		72 °C	3 Minutes
Denaturation	8	94 °C	1 Minute
Annealing		58 °C	1 Minute
Elongation		72 °C	3 Minutes
Denaturation	5	94 °C	1 Minute
Annealing		56 °C	1 Minute
Elongation		72 °C	3 Minutes
Final Elongation	1	72 °C	10 Minute
Optional Step		4 °C	Hold

The amplicons obtained have been subsequently prepared for the CE analysis onto the ABI 3500 Geneic Analyser (Applied Biosystems). Briefly, 2 µL of each PCR product has been mixed with 0.5 µL of GeneScan™ 500 LIZ™ dye (Applied Biosystems) and 7.5 µL of Formamide, for a total volume of 10 µL. The mixture has

been denatured for 3 minutes at 95 °C and immediately placed on ice for 2 minutes. Eventually, PCR products have been loaded on the instrument and the fragment analysis has been performed. Gene Mapper v5 has been used to analyse samples and to determine the dimension of the HRE. Due to the low resolution and the impossibility to precisely count the repeat units using this protocol, positive samples have been validated also with the AmplideX PCR/CE C9orf72 Kit (Asuragen) following the manufacturer's instructions. This protocol allows the detection and the count of the $(4G2C)_n$ repeats up to 145 units.

15. Multiplex Ligation-dependent Probe Assay (MLPA)

The presence of exon deletion and/or duplication in the GAA gene has been investigated via MLPA, using the SALSA MLPA P453 GAA probe mix (MRC Holland; for details of the protocol, see http://www.mlpa.com) following manufacturer's instructions. Amplified fragments have been separated by capillary electrophoresis using the ABI3500 Genetic Analyser and results have been analysed using the Coffalyser.Net software (MRC Holland). Obtained fluorescent data underwent intra-sample and inter-sample normalization. Exons Dosage Quotient (DQ) has been calculated considering normal DQ between 0.80 and 1.20. A DQ value <0.65 indicates the presence of a deletion, while a DQ value >1.3 a duplication. Possible exonic deletions or duplications were confirmed by real-time PCR analysis.

16. Peripheral blood mononuclear cells (PBMCs) isolation

PBMCs were isolated from fresh peripheral venous blood through a density gradient centrifugation using the Histopaque®-1077 (density = 1,077; Sigma-Aldrich). The gradient ensures the separation of cells during centrifugation and thus their easily collection. Briefly, a 1:1 ratio of Ficoll-Histopaque®-1077 and blood were centrifuged at 1800 rpm for 30 minutes at RT with maximum acceleration to allow phases' stratification and minimum deceleration to avoid the mixing among phases (Figure 13). The translucent white ring between the plasma and the Histopaque®-1077 represents PBMCs. Cells were recovered in a new tube, washed once with 1X Phosphate Buffer Saline Solution (PBS; Sigma-Aldrich), and centrifuged at 1500 rpm for 10 minutes at RT. Cells were then conserved at -80 °C for further applications.

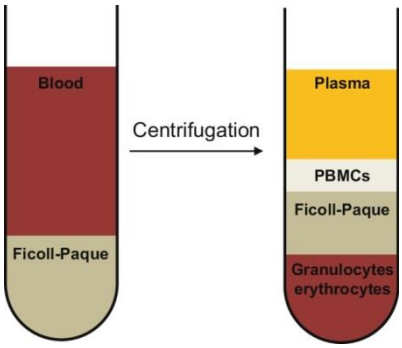

Figure 14| Ficoll-Paque™ gradient enrichment of PBMCs. The left tube represents Ficoll-Paque™ and blood before centrifugation, and the right tube represents separated phases.[6]

17. Extraction of RNA and quality control

RNA extraction has been made starting from previously isolated PBMCs using the guanidinium thiocyanate-phenol-chloroform method (TRIzol, Sigma-Aldrich). TRIzol works by maintaining RNA integrity during tissue homogenization, while at the same time disrupting and breaking down cells and cell components. As for DNA, the quantitative assessment of RNA amount has been made through spectrophotometric assay (Nanodrop 1000, Thermo Fisher Scientific). The purity of the obtained value (260/280 e A260/230 absorbance ratio) has been considered acceptable between 1.8-2.2. RNA has been stored at -80 °C until further usage.

18. Reverse transcription

Reverse transcription has been done to generate complementary DNA (cDNA) starting from an RNA template, using the iScript cDNA Synthesis Kit (Bio-Rad). 1000 ng of RNA previously quantified has been used as starting material in a final volume of 20 µL in order to get the final concentration of 50 ng/ µL. The reverse transcription reaction has been set up as follow:

Reagent	Time
5X iScript reaction Mix	4 µL
iScript Reverse Transcriptase	1 µL
Nuclease-free Water	Variable
RNA 1000 ng	Variable
Total Volume	20 µL

The 5X iScript reaction Mix contains all the components required for cDNA synthesis, including buffer, dNTPs, oligo(dT) and random primers, and a ribonuclease inhibitor.

The oligo(dT) are constructed to anneal specifically at the Poly(A) tails of the RNA template, while the random primers are made to anneal at random sites of the RNA template, to efficiently reverse transcribe all the messenger RNAs.
The program used on the thermocycler is described below:

	Number of cycles	Temperature	Time
Priming	1	25 °C	5 Minutes
Reverse Transcription (RT)	1	46 °C	20 Minutes
RT inactivation	1	95 °C	1 Minute
Optional Step	1	4 °C	hold

19. Real-Time PCR

Real-Time PCR has been used to confirm results found through MLPA. Particularly, per each sample presenting a deletion and/or a duplication of specific exons, Real-Time PCR has been performed on the exons of interest and on the flanking exons. The PCR reaction has been carried out on gDNA by using the iQ SYBR® Green Supermix (BioRad), in a final volume of 20 µL, according to the following proportions:

Reagent	Volume
iQ SYBR® Green Supermix	10 µL
Primer forward (10 µM)	1 µL
Primer reverse (10 µM)	1 µL
gDNA (10 ng/µL)	3 µL
H_2O	up to 20 µL

PCR cycling conditions

	Number of cycles	Temperature	Time
Initial denaturation	1	95 °C	4 Minutes
Denaturation	40	95 °C	15 seconds
Annealing* / Elongation		T_m	60 seconds

*The optimal annealing temperature is dependent upon the primer sequences and is usually 2-5 °C below the lower Tm of the pair.

Two housekeeping genes, the GAPDH (Glyceraldehyde-3-Phosphate Dehydrogenase) and the B2M (Beta-2-Microglobulin) were coamplified with each individual as internal standards. Individuals that carry normal copies of the GAA gene were used as controls. Specific primers using www.bioinformatics.nl/primer3plus were designed (primer sequences available upon request). Gene dosage was performed using the 2–ΔΔCT method according to Livak and Schmittgen.[302]
Real-time PCR has been also performed to assess transcript levels of GAA. The protocol adopted was the same, however it was performed on cDNA instead of gDNA.

20. 3D-modeling

To assess the influence of three single nucleotide polymorphisms (SNPs) on the structure and on the stability of the *GAA* enzyme, a three-dimensional (3D) model using Pcons (http://pcons.net/)[303] of the GAA enzyme has been created, in collaboration with the laboratory of Pathology and Immunology of the University of Pavia. The model has been created by superimposing the three models available from the PDB database (3llpA, 3lpoA, 3w37) to the model created using Pcons. Mutations have been subsequently added to the model to study the stability, the interaction surface and possible changes in secondary structure. DUET (http://biosig.unimelb.edu.au/duet/),[304] mCSM (http://biosig.unimelb.edu.au/mcsm/)[305] and SDM (http://marid.bioc.cam.ac.uk/sdm2)[306] tools were used to investigate the stability, InterPro (https://www.ebi.ac.uk/interpro/entry/InterPro/IPR003483/)[307] was used to assess surface interactions and the DSSP and Stride plugins for PyMol (https://pymol.org/2/; Schrödinger) were used to examine changes in secondary structure.

Results

21. AMYOTROPHIC LATERAL SCLEROSIS

21.1 Phenotypic characterization of the ALS cohort

The clinical dataset was complete for 231 ALS cases except for the progression rate that has been recorded for 193 patients (Table 6). Our cohort comprises 56% (n = 129) of males and 78% (n = 180) of spinal onset cases. The percentage of cases that present also neurological involvement was 19,5% (n = 45), and particularly the 7,4% (n = 17) present an ALS-FTD spectrum, while the 10% (n = 23) have cognitive impairment (ALSci) and the 2,1% (n = 5) have behavioural changes (ALSbi). Phenotypically, 61% (n = 140) of cases have a classical phenotype, 14% (n = 33) has a pure lower motor neuron (PLMN) disease, 7% (n = 16) has upper motor neuron (PUMN) phenotype, 10% (n = 22) of cases is bulbar, 3% (n = 7) is flail arm and another 3% have a pyramidal phenotype, 1% (n = 3) is flail leg, while the remaining three cases have a mix phenotype. 34% (n = 65) of cases was characterized by a slow progression rate (ALSFR < 0.5), followed by 40% (n = 77) of cases with an intermediate phenotype (ALSFR comprised between 0.5 and 1), while the remaining 26% (N = 51) of cases have a fast (ALSFR > 1) progression rate. The average age of onset is 63,5 years (SD = 11,67 years). A total of 39 patients resulted to have a positive family history for neurological pathologies. In particular, 10 patients for stroke, 9 patients for dementia or AD, 8 patients for PD, and 4 patients for ALS or other motor neuron diseases.

Clinical features	number	percent (%)
Sex:		
Male	129	56%
Female	102	44%
Age of onset (mean)	63,5 (SD = 11,67 years)	
Site of onset:		
Spinal	180	78%
Bulbar	40	17%
Spinal+Bulbar	11	5%
NP involvement:		
None		
ALS-FTD	17	7,40%
ALSci	23	10%
ALSbi	5	2,10%

Clinical features	number	percent (%)
Phenotypes:		
Classic	140	61%
Bulbar	22	10%
Pyramidal	7	3%
PUMN	16	7%
PLMN	33	14%
Flail arm	7	3%
Flail leg	3	1%
Miscellaneous	3	1%
Rate of progression:		
Fast	51	26%
Intermediate	77	40%
Slow	65	34%

Table 6| clinical features of 231 out of 248 ALS cases. Number of cases and relative percentage has shown. For the age of onset, mean age at diagnosis and standard deviation has been calculated.

Statistical analyses were performed to identify any association between clinical features. According to Welch's t-test, cases with bulbar onset were more likely to have a later age of onset (p=0.0003, figure 14A). χ2 analysis revealed significant association between phenotypes and site of onset (p=0.0001, figure 14B), where spinal onset resulted more prevalent in all phenotypes except the bulbar one, suggesting that the spinal site of onset can lead to more diverse phenotypes. We found also a significant association (p=0,0372, figure 14C) between neuropsychological involvement and age of onset when stratified into three groups: juvenile-onset (age < 45), classic-onset (45 ≤ age ≤ 65) and late-onset (age > 56). Juvenile-onset patients did not show any neuropsychological involvement, while in late-onset patients a cognitive impairment is more prevalent with respect the other neuropsychological outcomes. Eventually, a significant association (p=0,0123, figure 14D) between rate of progression and ALS phenotypes has been observed, where PUMN shows a more prevalent slower phenotype, while mixed phenotypes have a faster rate of progression.

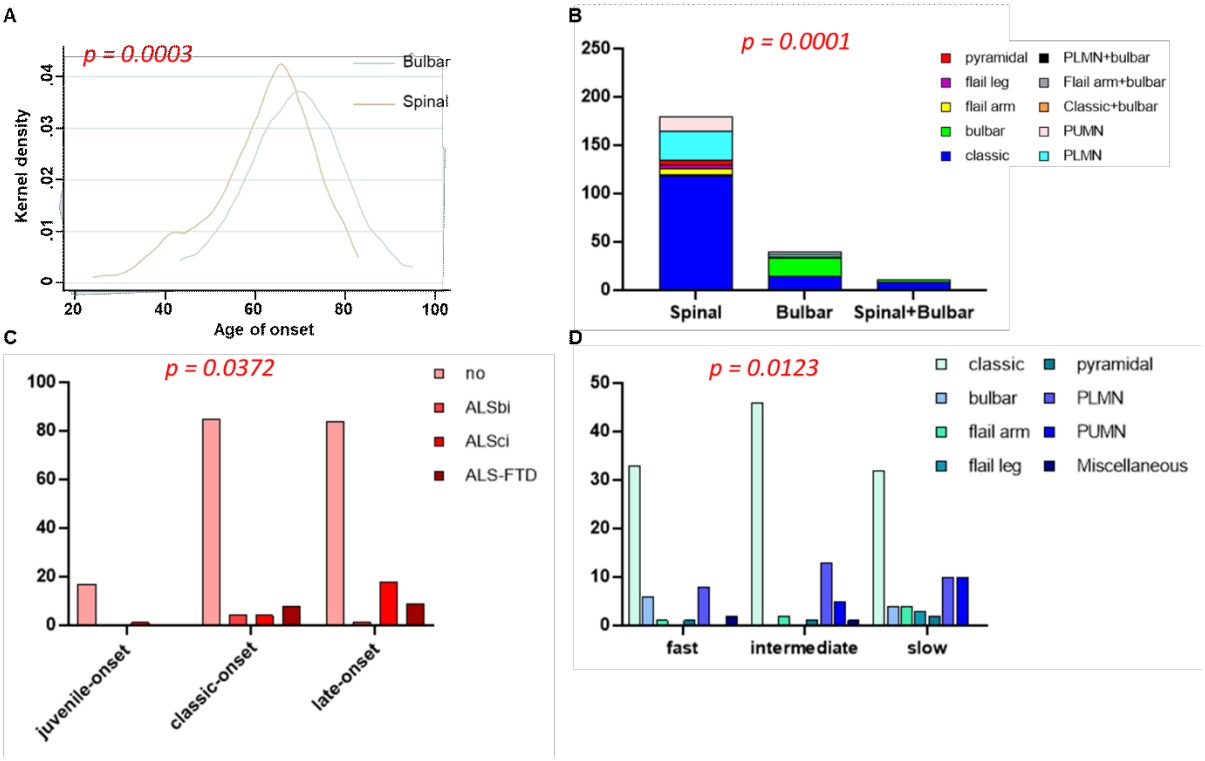

Figure 15| Statistical analyses of clinical variables for 248 ALS cases. A) shows the significant association between age of onset and site of onset; B) and D) show significant differences in site of onset and rate of progression between different phenotypes respectively; C) show differences in neuropsychological involvement based on the age of onset.

21.2 Panel 1 (97 genes for 60 patients)

Through this gene panel, and the independent *C9ORF72* mutational analysis, we found a total of 52 variants in 36 patients (60%) out of 60 patients (Figure 15A). According to the ACMG criteria, nine variants (17,31%) were classified as pathogenic, ten variants (19,23%) were classified as likely pathogenic and the remaining thirty-three variants (63,46%) were classified as VUS (Figure 15B).

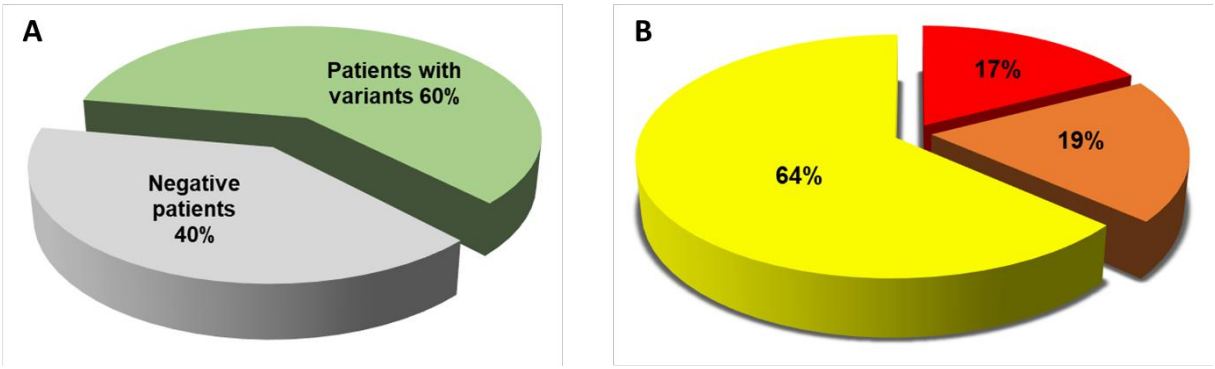

Figure 16| A) Pie chart shows the percentage of positive and negative patients in panel 1; B) Pie charts shows the percentage of pathogenic (red), likely pathogenic (orange) and VUS (yellow) variants.

21.2.1 Pathogenic and Likely pathogenic variants

Among the pathogenic variants, six were expansions in the *C9ORF72* gene, two were heterozygous mutations in the *SPG7* gene while one was found also in heterozygous state in the *SPG11* gene; among the likely pathogenic variants, two were found in the *VCP* gene, while one variant was found in the *TARDBP*, *SOD1*, *FUS*, *CYP27A1*, *SPG7*, *PINK1*, *POLG* and *MNF2* genes respectively (Table 7). All variants were heterozygous. In particular, the likely pathogenic variants in *TARDBP* and *POLG* were found in two patients who carry already mentioned pathogenic mutations, while the likely pathogenic variant in *SPG7* was found in the patient carrying the likely pathogenic mutation in *FUS*. In total, twelve out of the nineteen variants were in genes associated with ALS (*C9ORF72*, *TARDBP*, *SOD1*, *FUS*, *VCP* and *CYP27A1*), four mutations were in genes associated with HSP (*SPG7* and *SPG11*), two were in PD-linked genes (*POLG* and *PINK1*), and only one variant was in a gene (*MNF2*) linked to other neuromuscular diseases, in particular to CMT type 2A2A (MIM:609260) and type 2A2B (MIM:617087), and to Hereditary motor and sensory neuropathy type VIA (MIM:601152). These diseases have all an AD pattern of inheritance.

Gene	Variant cDNA	Variant protein	MAF*	dbSNP	ACMG Verdict	N° of pts	Patient	Other variants
C9ORF72	**-**	**-**	**-**	**-**	-	6	P005	***FUS:*** c.221G>C p.Gly74Ala (VUS)
							P015	
							P027	***SETX*****:** c.3056C>A p.Ser1019Tyr (VUS); ***SARM1*****:** c.893C>A p.Pro298Gln (VUS)
							P030	
							P045	***SARM1:*** c.893C>A p.Pro298Gln (VUS)
							P059	
SPG7	c.1468G>T	p.Glu490*	0.0% 0.0%	rs769688149	P	1	P004	--
SPG7	c.1529C>T	p.Ala510Val	0.3% 0.2%	rs61755320	P	1	P033	***POLG***: c.926G>T p.Arg309Leu (L. Pathogenic); ***HNRNPA1***: c.883G>A p.Gly295Arg (VUS)
SPG11	c.5986dupT	p.Cys1996fs	0.0% 0.0%	rs312262775	P	1	P021	***TARDBP:*** c.1070G>A p.Gly357Asp (L. Pathogenic)
VCP	c.277C>T	p.Arg93Cys	0.0% 0.0%	rs1554669087	LP	1	P60	-
VCP	c.791C>G	p.Arg191Gln	0.0% 0.0%	rs121909334	LP	1	P018	***AP5Z1***: c.383_385delAGG p.Glu128del (VUS)
SOD1	c.442G>A	p.Gly148Ser	0.0% 0.0%	NA	LP	1	P002	-
FUS	c.1520G>A	p.Gly507Asp	0.0% 0.0%	rs267606831	LP	1	P051	***SPG7***: c.1796G>T p.Arg599Leu (L. Pathogenic)
CYP27A1	c.1183C>T	p.Arg395Cys	0.0% 0.0%	rs121908096	LP	1	P057	-
PINK1	c.802C>G	p.Leu268Val	0.0% 0.0%	rs372280083	LP	1	P001	**-**
MFN2	c.349C>A	p.Leu117Ile	0.0% 0.0%	NA	LP	1	P052	-

Table 7| List of pathogenic and likely pathogenic variants found in Panel 1. Gene name, nucleotide change, amino acidic change, frequency in the population databases, dbSNP ID, ACMG verdict, the n° of patients and second variants within the patients are shown. **P**: pathogenic; **LP**: likely pathogenic; **VUS**: variants of uncertain significance; **NA**: not available. *MAF (minor allele frequency) in GnomAD and 1000 genome databases.

21.2.2 Variants of uncertain significance

Among the 52 variants found in Panel 1, thirty-three variants (63,46%) were of uncertain significance according to the ACMG criteria (Figure 16). Fourteen variants out of thirty-three were found in genes causative for ALS (*FUS*, *HNRNPA1*, *SETX* and *SQSTM1*) or that confer susceptibility to ALS (*APEX*, *NEK1*, *PRPH*, *SARM1*, *SS18L1* and *TAF15*), while six variants were respectively found in genes associated to HSP (*AP5Z1*, *BSCL2*, *L1CAM*, *SPG21* and *ZFYVE26*) and to PD (*GIGYF2*, *POLG* and *SPR*). The remaining seven variants were present in genes (*ATP2B3*, *ATXN3*, *BAG3*, *DHTKD1*, *IGHMBP2* and *UBA1*) associated with other NDs, mostly neuromuscular diseases such as CMT type 2Q or 2S, Distal hereditary motor neuropathy type VI, Spinocerebellar ataxia, Machado-Joseph disease and Spinal muscular atrophy.

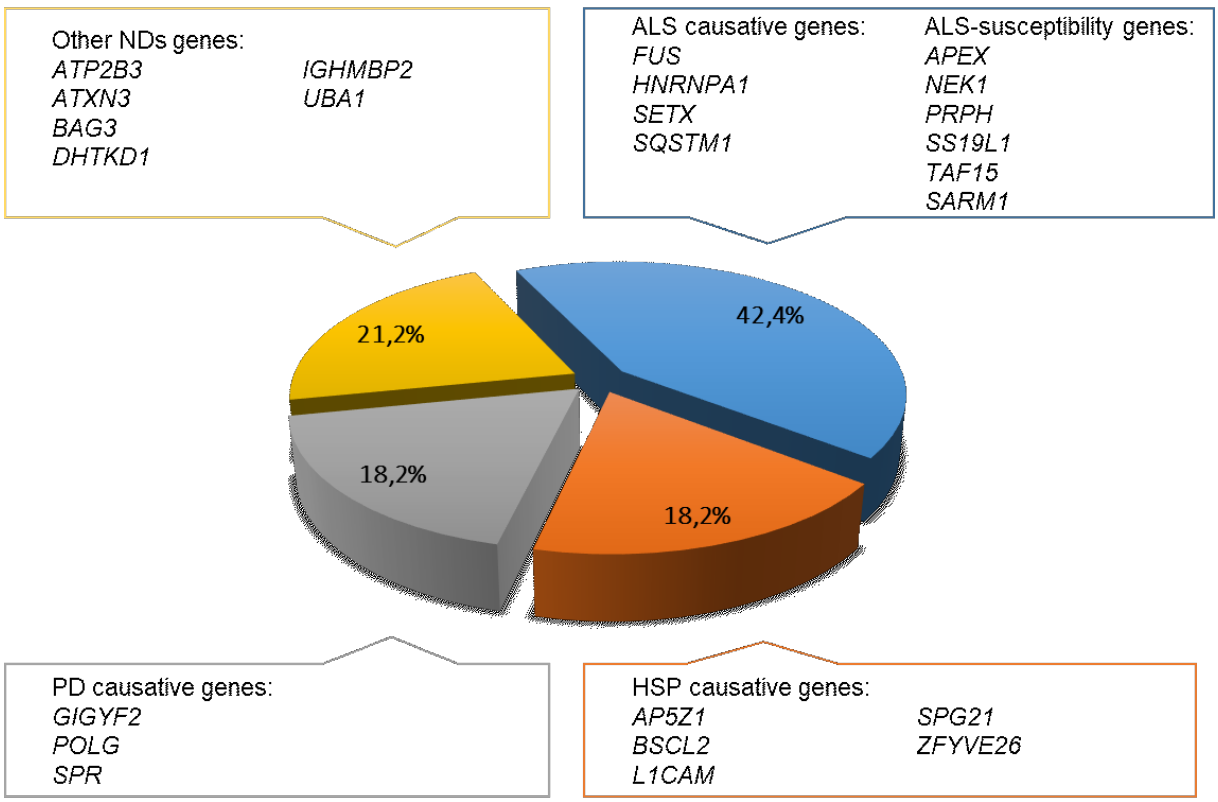

Figure 17| graphic categorization of the VUS variants found in Panel 1 based on the belonging group of disease: ALS (blue), HSP (orange), PD (grey), other NDs (yellow).

From a functional point of view, the majority of variants (n = 25) were missense, followed by variants in the splice-site region (n = 4) thus that may have effects on the transcript length, variants that alter the 5'-UTR region which create an anticipated start codon (n = 2) and by variants that disrupt the reading frame (n= 1) ending up in

a altered protein sequence, and that introduce a stop codon (n = 1) which results in the formation of a truncated protein (Figure 17).

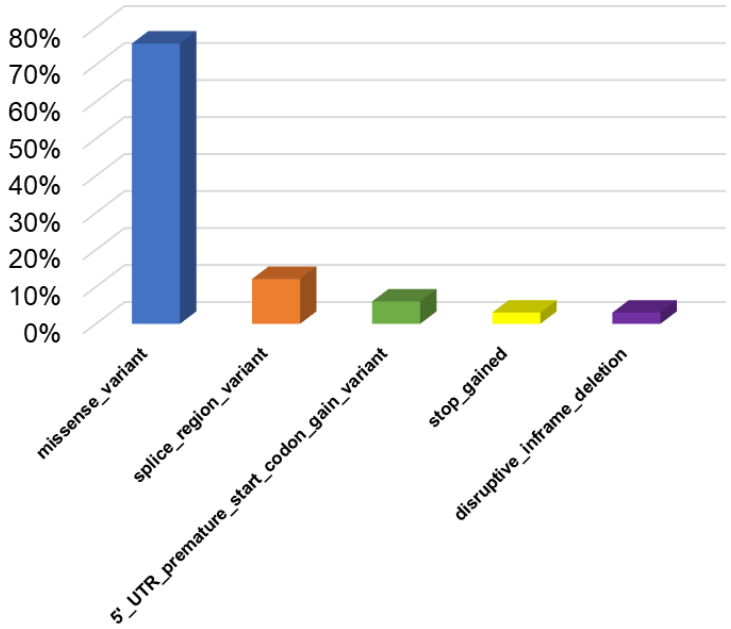

Figure 18| Functional characterization of the VUS variants found in Panel 1: missense variants (blue), splice-region variants (orange), 5'-UTR variants (green), stop-gain variants (yellow) and inframe variants (purple). X axis represents the relative percentage of the different types of VUS variants found.

21.3 Panel 2 (173 genes for 107 patients)

Through this gene panel, and the independent *C9ORF72* mutational analysis, we found a total of 189 variants in 90 patients (84,1%) out of 107 patients (Figure 18A). According to the ACMG criteria, eighteen variants (9,5%) were classified as pathogenic, nine variants (4,5%) were classified as likely pathogenic and the remaining one hundred and sixty-one variants (86%) were classified as VUS (Figure 18B).

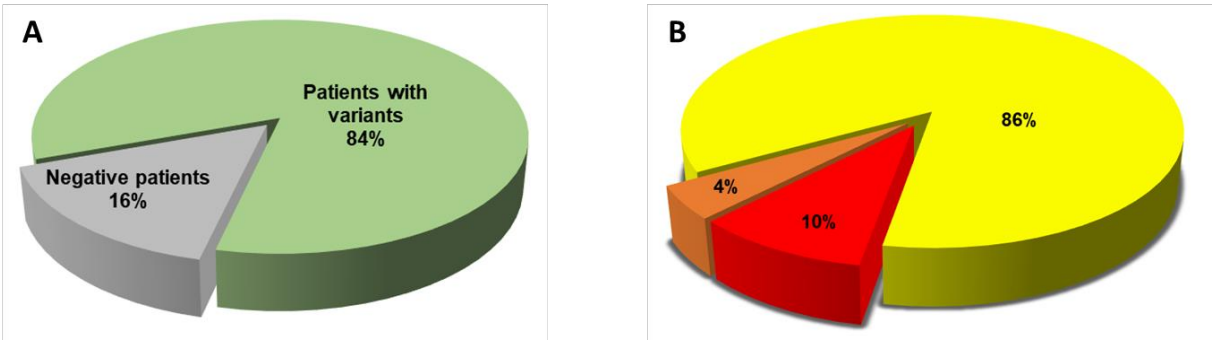

Figure 19| A) Pie chart shows the percentage of positive and negative patients in panel 2; B) Pie charts shows the percentage of pathogenic (red), likely pathogenic (orange) and VUS (yellow) variants.

21.3.1 Pathogenic and Likely pathogenic variants

Among the pathogenic variants, seven were expansions in the *C9ORF72* gene, two were heterozygous mutations in the *DCTN1* gene, two were found in the *HEXA* gene, one heterozygous variant was found in the *ALS2* gene, one mutation was in the SQSTM1 gene, one was present in the *SPG7* gene, one was in the *GBA* gene, and other thee were found respectively in *TREM2*, *LRRK2* and *PRKN*; the likely pathogenic variants were all heterozygous and were found in the *CYP27A1*, *KIF5A*, *GDAP1*, *HSPB8*, *GCH1*, *PINK1*, *POLG*, *PSEN1* and *HEXA* genes respectively (Table 8). The pathogenic *SPG7* variant was found in two patients, while the pathogenic variant in *TREM2* was found in the same patient carrying also one of the two pathogenic mutation in the *DCTN1* gene. In total, fourteen out of the twenty-seven variants were in genes associated with ALS (*C9ORF72*, *DCTN1, ALS2, SQSTM1, KIF5A*, *CYP27A1 and TREM2*), one variant was within a gene associated with HSP (*SPG7*), five were in PD-linked genes (*GBA*, *LRRK2*, *PRKN*, *PINK1* and *POLG*) and the remaining eight variants were in genes linked to other NDs (*GDAP1*, *HSPB8*, *GCH1*, *PSEN1* and *HEXA*). Except for *PSEN1*, which is a gene causative for AD, the other genes are associated to neuromuscular diseases with an AD inheritance

pattern, like CMT type 2L (MIM:608673), Distal hereditary motor neuropathy type IIA (MIM:158590) and CMT type 4A (MIM: 214400), or with diseases that may have either an AD and an AR pattern of inheritance, like Dystonia with or without hyper-phenylalaninemia (MIM:128230) and CMT type 2K (MIM:607831).

Gene	Variant cDNA	Variant protein	MAF*	dbSNP	ACMG Verdict	N° of pts	Patient	Other variants
ALS2	c.1447_1448delTC	p.Ser483fs	0.0% 0.0%	NA	P	1	P079	--
C9ORF72	**-**	**-**	**-**	**-**	P	7	P067	***SS18L1:*** c.605A>C p.Gln202Pro (VUS); ***TRIM28:*** c.1285C>A p.Pro429Thr (VUS)
							P088	--
							P148	***SMN1***: c.859G>C p.Gly287Arg (VUS); ***SYNE1***: c.4093C>T p.Arg1365Cys (VUS)
							P158	**ARHGEF28**: c.2075G>A p.Cys692Tyr (VUS); **ZHX2**: c.331G>A p.Val111Met (VUS)
							P163	**ABCA7**: c.2146T>C p.Trp716Arg (VUS); **UNC13A**: c.3832-5_3832-3delCTT p.?? (VUS)
							P165	***CYP27A1***: c.851A>G p.Lys284Arg (VUS); ***EIF4G1***: c.74C>Ap.Pro25Gln (VUS); ***GBE1***: c.293T>G p.Val98Gly (VUS); ***SMN1***: c.859G>C p.Gly287Arg (VUS)
							P167	***FGGY***: c.938G>A p.Trp313* (VUS)
CYP27A1	c.1435C>G	p.Arg479Gly	0.0% 0.0%	rs72551322	LP	1	P105	***TAF1***: c.1118T>C p.Met373Thr (VUS)
DCTN1	c.279+1G>T	p. ??	NA 0.0%	NA	P	1	P071	***TREM2***: c.97C>T p.Gln33* (P); ***SOD1***: c.272A>C p.Asp91Ala (VUS)
DCTN1	c.3609+1G>T	p. ??	0.0% 0.0%	NA	P	1	P089	-
KIF5A	c.2341A>G	p.Lys781Glu	NA 0.0%	NA	LP	1	P130	***ATP13A2***: c.3083+6T>G p.?? (VUS); ***BSCL2***: c.1286C>T p.Ala429Val (VUS); ***GIGYF2***: c.354A>T p.Arg118Ser (VUS)
SQSTM1	c.240C>G	p.Asp80Glu	0.0% 0.0%	rs148366738	P	1	P119	***ATP13A2***: c.2252-8A>T p.?? (VUS)
GDAP1	c.620A>C	p.Lys207Thr	0.0% 0.0%	rs140384868	LP	1	P070	***SYNE1***: c.11680G>T p.Asp3894Tyr (VUS)
HSPB8	c.421A>G	p.Lys141Glu	NA 0.0%	rs104894351	LP	1	P108	BRWD3: c.2045-8delT p.?? (VUS); ***FEZF2***: c.1120+6G>A p.?? (VUS); ***SYNJ1***:

								c.51_62delTGGC GGCGGCTG p.Gly18_Cys21del (VUS)
SPG7	c.1529C>T	p.Ala510Val	0.3% 0.2%	rs61755320	P	2	P144	***UBQLN2***: c.809G>A p.Arg270His (VUS); VPS13A: c.7400G>C p.Gly2467Ala (VUS)
							P153	***PRKN***: c.1310C>T p.Pro437Leu (VUS)
GCH1	c.242C>T	p.Ser81Leu	NA 0.0%	NA	LP	1	P096	***FIG4***: c.122T>C p.Ile41Thr (VUS); ***SORL1***: c.6525C>G p.Ser2175Arg (VUS); ***VPS13C***: c.950C>T p.Thr317Met (VUS)
GBA	c.1448T>C	p.Leu483Pro	0.1% 0.3%	rs421016	P	1	P155	**SHC2**: c.691C>G p.Leu231Val (VUS)
LRRK2	c.3632delT	p.Leu1211fs	NA 0.0%	NA	P	1	P160	-
PRKN	c.823C>T	p.Arg275Trp	0.2% 0.0%	rs34424986	P	1	P141	-
PINK1	c.803T>C	p.Leu268Pro	NA 0.0%	NA	LP	1	P142	**ERBB4**: c.1441A>G p.Ile481Val (VUS); ***TRPV4***: c.394C>G p.Pro132Ala (VUS)
POLG	c.1399G>A	p.Ala467Thr	0.1% 0.0%	rs113994095	LP	1	P132	**ARHGEF28**: c.3089C>G p.Ala1030Gly (VUS); ***DNAJC6***: c.2250A>C p.Lys750Asn (VUS); ***GBA***: c.308-8G>C p.?? (VUS)
PSEN1	c.376G>A	p.Gly126Ser	NA 0.0%	NA	LP	1	P135	**SPG11**: c.6478-4delA p.?? (VUS)
HEXA	c.1496G>A	p.Arg499His	0.0% 0.0%	rs121907956	LP	1	P091	**PINK1**: c.587C>T p.Pro196Leu (VUS); ***SIGMAR1***: c.217G>A p.Val73Met (VUS)
HEXA	c.1073+1G>A	p. ??	0.0% 0.0%	rs76173977	P	1	P117	**FUS**: c.1056C>G p.Asp352Glu (VUS)
HEXA	c.533G>A	p.Arg178His	0.0% 0.0%	rs28941770	P	1	P124	**AP5Z1**: c.1132G>A p.Gly378Arg (VUS); ***AP5Z1***: c.1708-6C>T p.?? (VUS); ***FBXO7***: c.65C>T p.Thr22Met (VUS); ***SYNE1***: c.6880G>A p.Val2294Ile (VUS)

Table 8| List of pathogenic and likely pathogenic variants found in Panel 1. Gene name, nucleotide change, amino acidic change, frequency in the population databases, dbSNP ID, ACMG verdict, the n° of patients and second variants within the patients are shown. **P**: pathogenic; **LP**: likely pathogenic; **VUS**: variants of uncertain significance; **NA**: not available. *MAF (minor allele frequency) in GnomAD and 1000 genome databases

21.3.2 Variants of uncertain significance

Among the 189 variants found in Panel 2, one hundred and sixty-two variants (86%) were of uncertain significance according to the ACMG criteria (Figure 19). Forty-two variants were found in genes causative for ALS (*ALS2*, *ERBB4*, *FIG4*, *FUS, GLE1, HNRNPA3*, *KIF5A, OPTN*, *SETX*, *SIGMAR1*, *SOD1*, *SPG11*, *UBQLN* and *VCP*) or

that confer susceptibility to ALS (*CRYM*, *CYP27A2*, *DPP6*, *FGGY*, *NEK1*, *PRPH*, *SARM1*, *SS18L1* and *UNC13A*), twelve variants were found in genes associated to HSP (*AP5Z1*, *BSCL2*, *KIF1A*, *SPG7*, *SPG21* and *ZFYVE26*) and thirty-one were found in PD-genes (*ADORA1, ATP13A2, DNAJC6, EIF4G1, FBXO7, GBA, LRRK2, PRKN, PINK1, PLA2G6, POLG, SYNJ1, TAF1 and VPS13C*). Almost half of the VUS (n = 77) were instead present in genes associated with other NDs. Also in this case, several variants are in genes associated in other neuromuscular disorders, such as CMT type 2A2A (MIM:609260), 2A2B (MIM:617087), 2D (MIM:601472), 2N (MIM:613287) and 2Q (MIM:615025), Distal hereditary motor neuropathy type IIC (MIM: 613376) and VA (MIM:600794), and Spinal and bulbar muscular atrophy of Kennedy (MIM:313200), but also with other NDs that may share clinical manifestations with ALS such as Neurodegeneration with brain iron accumulation (MIM:234200) and Spinocerebellar ataxia 1 (MIM:164400).

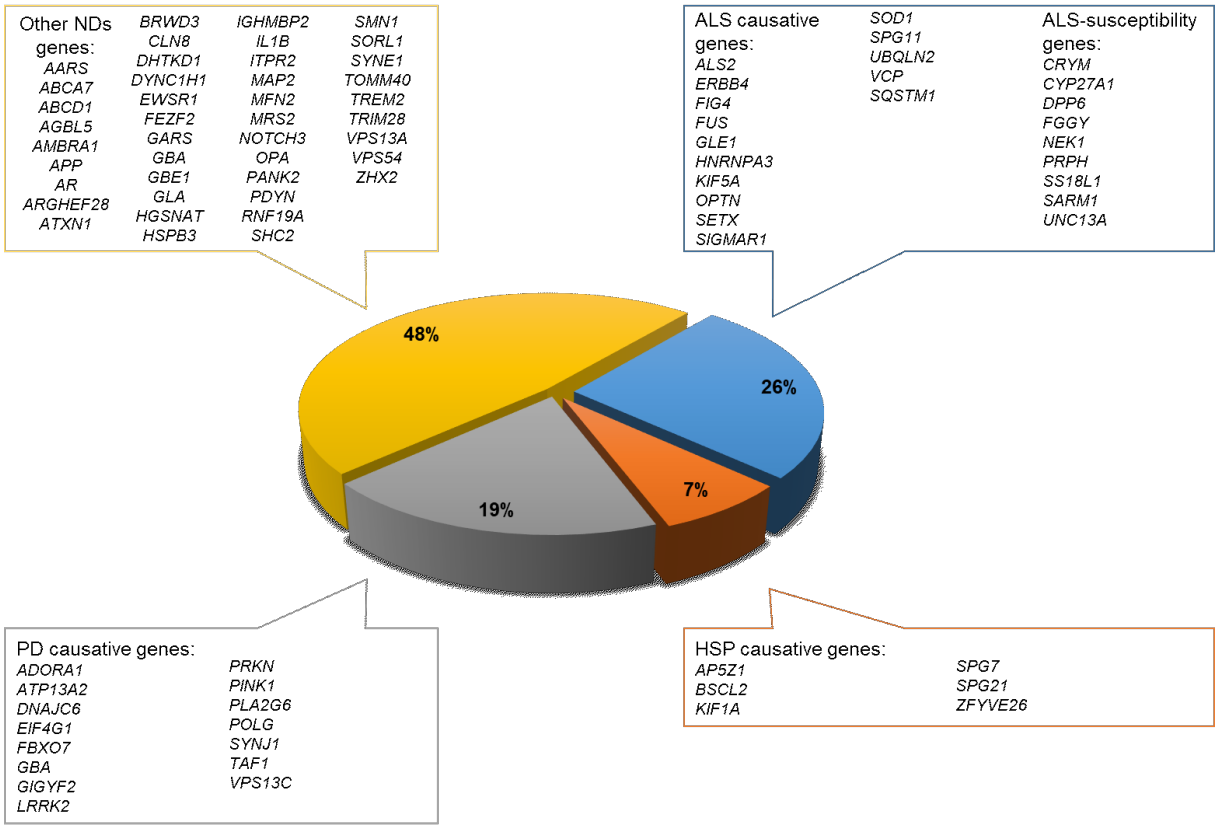

Figure 20| graphic categorization of the VUS variants found in Panel 1 based on the belonging group of disease: ALS (blue), HSP (orange), PD (grey), other NDs (yellow).

From a functional point of view, the majority of variants (n = 127) were missense as we found in Panel 1, followed by variants in the splice-site region (n = 20), variants that alter the 5'-UTR region (n = 6), variants that disrupt the reading frame (n= 3) or

small inframe indels (n = 3), and by variants that introduce a stop codon (n = 1) or that disrupt the initial ATG codon (n = 2) (Figure 20).

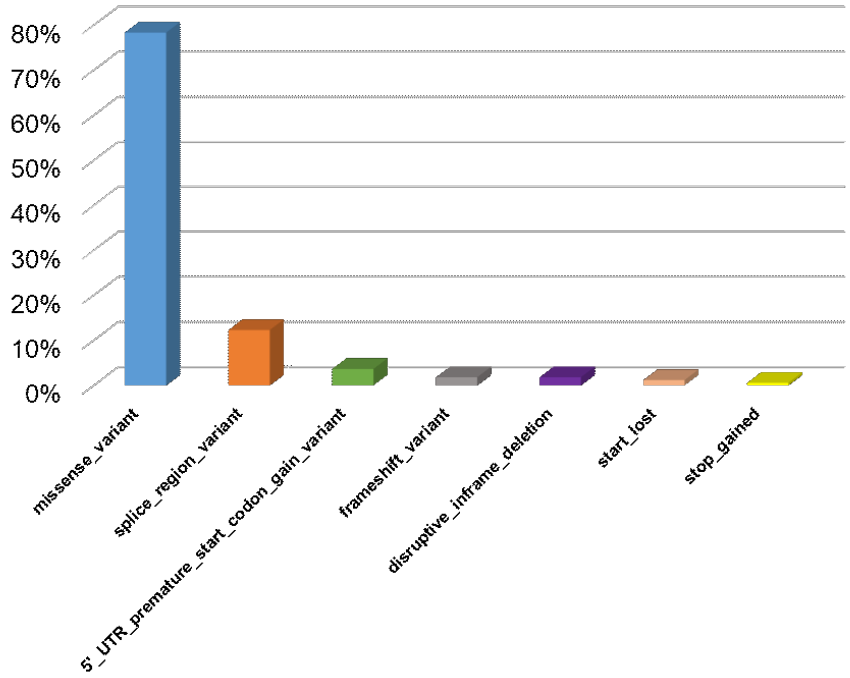

Figure 21| Functional characterization of the VUS variants found in Panel 2: missense variants (blue), splice-region variants (orange), 5'-UTR variants (green), frameshift variant (grey), inframe variants (purple), start loss variants (pink) and stop-gain variants (yellow). X axis represents the relative percentage of the different types of VUS variants found.

21.3.3 KIF5A: Case-report

Among the likely pathogenic in ALS genes, we found the c.2341A>G (p.Lys781Glu) missense variant in the exon 21 of the *KIF5A* gene (Figure 21). This gene encodes for the heavy chain of Kinesin-1, which is exclusively expressed in neurons and resides in cell body, dendrites and axon.[308] KIF5A is composed of a C-terminal globular tail involved in the cargo binding, a stalk domain which is an α-helical coiled-coil region involved in heavy chain dimerization, and of a N-terminal motor domain which is responsible for the motor activity along microtubules.[309] This gene has been very recently associated to ALS[142] and we found this mutation in a patient with a very peculiar clinical presentation, characterized by atypical motor syndrome and slow progression. The patient developed, over a period of at least four years, a slowly progressive motor syndrome, mainly involving proximal segments of limbs in an asymmetrical way. The patient presented signs of upper and lower motor neuron involvement, with the first ones more marked at lower limbs (pyramidal abnormalities and ankle clone) and the second ones at upper limbs (muscle hypotrophy, fasciculations).[310] The clinical presentation of the patient was in line with what

already reported in *KIF5A*-linked ALS patients in which an increased disease duration was described.[311]

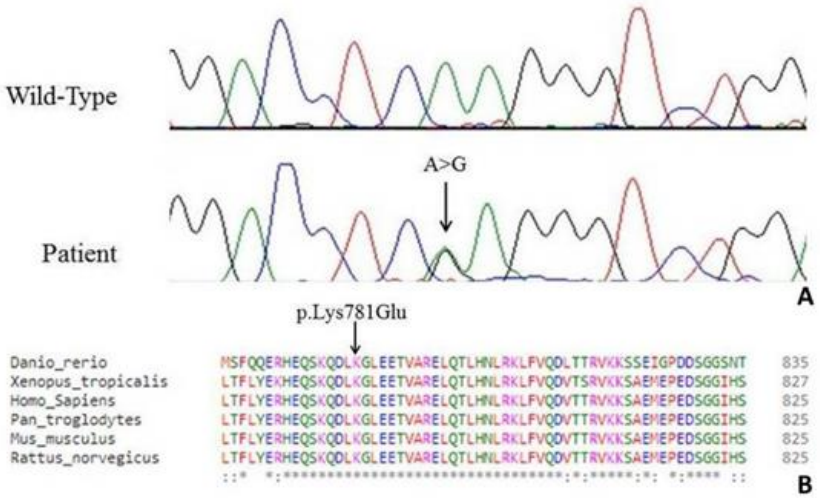

Figure 22| A) Electropherogram showing the novel heterozygote variant c.2341A>G in the exon 21 in the patient compared to a control sequence. B) Partial alignment of KIF5A from different species showing the residue Lys781 located within the terminal region of the stalk domain and highly evolutionary conserved.

The c.2341A>G nucleotide substitution is classified as likely pathogenic according to the ACMG criteria and was predicted as damaging by the major *in silico* prediction webtools (Poly-phen2, Mutation Taster and SIFT). RaptorX (http://raptorx.uchicago.edu/) software was used for molecular modelling prediction of the 3D structure both of the wildtype and of the mutated protein. We queried only the sequence of the domain where our mutation was placed, such as the stalk domain, from amino acid 331 to 906. The 3D wildtype structure predicted (p-value 7.14 × 10−5) revealed that the stalk domain is mainly formed by coiled alpha-helices structures (Figure 22A). The same was done to see how the 3D structure changed adding the Lys781Glu mutation (Figure 22B). As shown, the 3D structure of the stalk domain containing the mutation is highly altered compared to the wildtype (p-value 7.98 × 10−5), reinforcing the pathogenic role of our mutation.

A

B

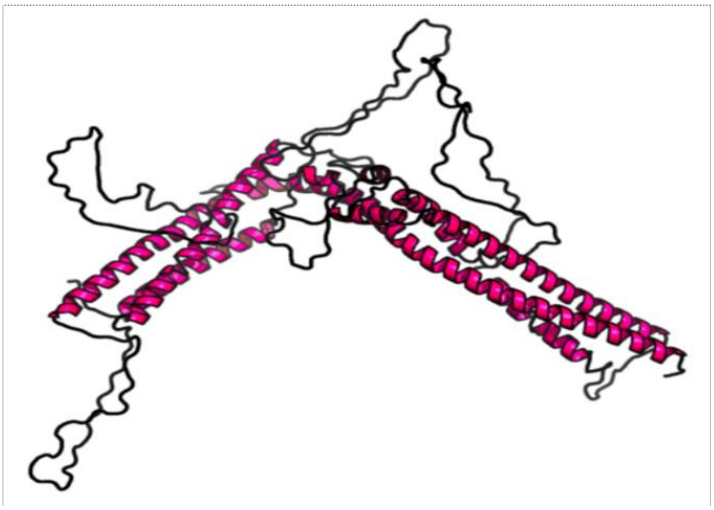

Figure 23| Image of the predicted 3D structure of the stalk domain of protein *KIF5A*. A) Wild-type structure of the stalk domain; B) Mutated structure of the stalk domain.

The stalk domain is characterized by a coiled coil region, where two or more alpha-helices are coiled together. In a coiled coil region, the hydrophobic amino acids are involved in the stability of the coiled structure, while charged amino acids are usually exposed to the surface, in contact with water molecules or involved in salt bridges to stabilize the 3D structure and allow the dimerization with other proteins. The Lysine 781 variation is a positively charged polar amino acid that changes into a Glutamate, which instead is a negatively charged polar amino acid. Moreover, based on Genomic position to 3D data (G23D, https://www.sheba-cancer.org.il/cgi-bin/variants/G23D.cgi) software, these two amino acids present a different spatial orientation that, together with the different electric charge, may alter both the interaction among alpha-helices in the coiled coil domain and the dimerization of the heavy chain (Figure 23A, 23B and 123C).

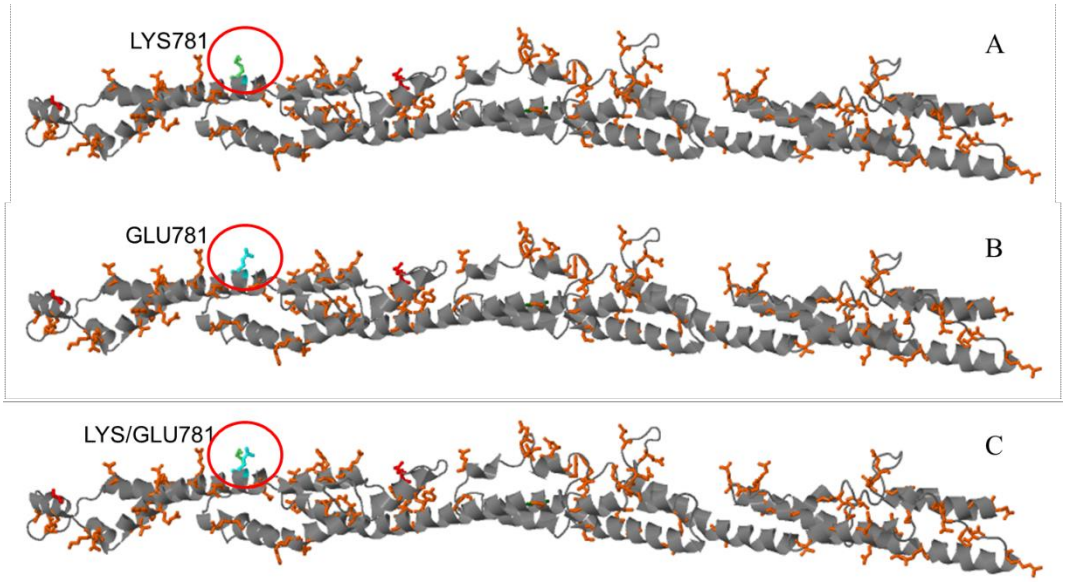

Figure 24| Image of the coiled coil region from amino acid 331 to amino acid 906 of the *KIF5A* protein and the spatial visualization of the wild-type and mutant amino acids at targeted position 781. A) Spatial visualization of the Lysine at position 781; B) Spatial visualization of the Glutamate at position 781; C) Spatial visualization of both the Lysine and the Glutamate at position 781.

The p.Lys781Glu variant falls in the terminal region of the stalk domain, outside from the motor domain which is a hot spot for HSP pathogenic mutations as well as from the cargo binding domain in which loss of function changes involved in ALS phenotype are located. Few other mutations outside of motor and cargo domains were described, falling in the stalk domains or in the neck region and they are associated with a HSP phenotype.[312,313] Therefore, our patient represents the first case, albeit atypical, of ALS phenotype associated to a stalk domain mutation. How mutations in the same gene affect transport activity either selectively in upper motor neurons and in peripheral axons or simultaneously in both central and peripheral motor pathways remain unclear.

In any case, whatever the underlying mechanism, our study confirmed that *KIF5A* mutations may be linked to slowly progressive and early-onset ALS phenotypes, even if patient's phenotype seems to suggest a more complex overlapping motor syndrome, as well as the importance to perform wide genetic screening also in sporadic ALS cases to help clinicians to unravel challenging clinical cases.

21.4 Panel 3 (277 genes for 81 patients)

Through this gene panel, and the independent *C9ORF72* mutational analysis, we found a total of 182 variants in 75 patients (92,6%) out of 81 patients (Figure 24A). According to the ACMG criteria, eleven variants (6,6%) were classified as pathogenic, twelve variants (6%) were classified as likely pathogenic and the remaining one hundred and fifty-nine variants (87,4%) were classified as VUS (Figure 24B).

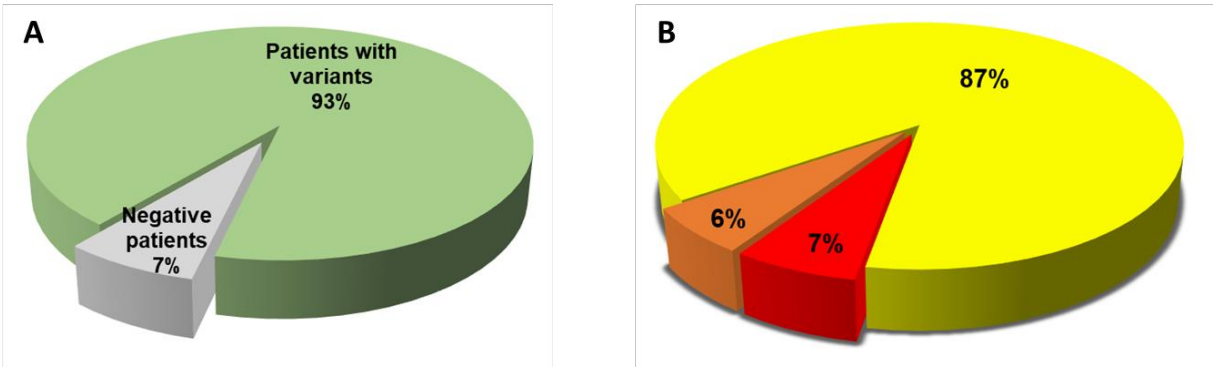

Figure 25| A) Pie chart shows the percentage of positive and negative patients in panel 3; B) Pie charts shows the percentage of pathogenic (red), likely pathogenic (orange) and VUS (yellow) variants.

21.4.1 Pathogenic and Likely pathogenic variants

Among the pathogenic variants, three were expansions in the *C9ORF72* gene, while the other eight were found in heterozygous state in the *SOD1*, *TARDBP*, *OPTN*, *FIG4*, *CYP2U1*, *GBA*, *PRKN* and *GLB1* genes; concerning the likely pathogenic variants, three were found in the *GLB1,* in particular two were contiguous within the same patient, while the others were found in the *NEK1*, *TARDBP*, *ALDH18A1*, *SPG7*, *NPC1*, *GBA*, *CSF1R*, *RNASEH2B* and *HEXA* genes respectively (Table 9). The likely pathogenic variants within *CSF1R* and *RNASEH2B* were found as additional variants in the patients carrying the pathogenic *GBA* and likely pathogenic *SPG7* variants respectively, as well as the pathogenic variant in the *PRKN* gene which was found in one patients with the (C2G4) repeat expansion in the *C9ORF72* gene. In total, nine out of the twenty-three variants were in genes associated with ALS (*C9ORF72, SOD1*, *TARDBP*, *OPTN*, *FIG4* and *NEK1*), three variants were within genes associated with HSP (*ALDH18A1*, *CYP2U1* and *SPG7*), six were in PD-linked genes (*GBA*, *PRKN* and *CSF1R*) and the remaining five variants were in genes linked to other NDs (*GLB1*, *RNASEH2B* and *HEXA*). These last three genes are respectively associated with GM1-gangliosidosis type I, II and III (MIM:230500;

MIM:230600; MIM:230650), with Aicardi-Goutieres syndrome 2 (MIM:610181), and with GM2-gangliosidosis (MIM:272800).

Gene	Variant cDNA	Variant protein	MAF*	dbSNP	ACMG Verdict	N° of pts	Patient	Other variants
C9ORF72	-	-	-	-	P	3	P177	***PLA2G6***: c.901C>T p.Arg301Cys (VUS); ***DDHD1***: c.1870C>A p.Pro624Thr (VUS);
							P182	***ABCA7***: c.1456C>G p.Pro486Ala (VUS); ***NOTCH3***: c.3758A>C p.Gln1253Pro (VUS); ***PRKN***: c.823C>T p.Arg275Trp (P); ***ZFYVE26***: c.1091C>T p.Pro364Leu (VUS); ***DNM2***: c.2305C>T p.Arg769Cys (VUS); ***NT5C2***: c.1677_1679delGGA p.Glu560del (VUS); ***SACS***: c.10577T>A p.Ile3526Lys (VUS)
							P184	--
OPTN	c.167-1G>A	p. ??	0.0% 0.0%	rs764745062	P	1	P197	--
SOD1	c.260A>G	p.Asn87Ser	0.0% 0.0%	rs11556620	P	1	168	***SYNE1:*** c.14573T>C p.Ile4858Thr (VUS); ***HSPD1:*** c.607-8T>C p.?? (VUS)
TARDBP	c.883G>A	p.Gly295Ser	0.0% 0.0%	rs80356723	P	1	P193	***ARHGEF28***: c.1658G>A p.Arg553His(VUS); ***KLC4***: c.1697C>A p.Thr566Asn (VUS); ***PDGFRB***: c.1531C>T p.Arg511Cys (VUS); ***SARM1***: c.893C>A p.Pro298Gln (VUS)
GBA	c.1448T>C	p.Leu483Pro	0.1% 0.3%	rs421016	P	3*	P187	***NOTCH3***: c.6089C>T p.Pro2030Leu (VUS); ***KIF5A***: c.1231C>T p.Pro411Ser
							P244	***CSF1R***: c.2512G>A p.Val838Ile (LP)
GLB1	c.1369C>T	p.Arg457*	0.0% 0.0%	rs72555359	P	1	P220	-
PRKN	c.823C>T	p.Arg275Trp	0.2% 0.0%	rs34424986	P	1	P184	***ABCA7***: c.1456C>G p.Pro486Ala (VUS); ***DNM2***: c.2305C>T p.Arg769Cys (VUS); ***SACS***: c.10577T>A p.Ile3526Lys (VUS); ***ZFYVE26***: c.1091C>T p.Pro364Leu (VUS); ***NOTCH3***: c.3758A>C p.Gln1253Pro (VUS); ***NT5C2***: c.1677_1679delGGA p.Glu560del (VUS)
CYP2U1	c.1630A>T	p.Arg544*	0.0% 0.0%	rs772136947	P	1	P229	***SARM1***: c.740A>G p.Lys247Arg (VUS)
ALDH18A1	c.1393G>A	p.Glu465Lys	0.0% 0.0%	rs757876226	LP	1	P180	****GBA***: c.1448T>C p.Leu483Pro (P); ***DNAJC6***: c.2811+4G>C p.?? (VUS); ***NAIP***:

								c.374G>A p.Gly125Glu (VUS); ***POLG***: c.3436C>T p.Arg1146Cys (VUS); ***TAF15*** c.1278_1304delCTAT GGTGGAGACAGAA GCAGCGGTGG p.Tyr427_Gly435del (VUS); ***AP4M1***: c.173A>G p.His58Arg (VUS); ***FLRT1:*** c.251C>T p.Thr84Ile (VUS)
TARDBP	c.881G>T	p.Gly294Val	0.0% 0.0%	rs80356721	LP	1	P238	***ERLIN1***: c.728G>A p.Arg243His (VUS); ***FGGY***: c.1372G>A p.Gly458Arg (VUS); ***SPG20***: c.1652C>T p.Thr551Ile (VUS)
NPC1	c.1618C>T	p.Pro540Ser	0.0% 0.0%	rs886044064	LP	1	P240	***DAO***: c.113G>A p.Arg38His (VUS); ***DNAJC13***: c.5159C>T p.Ser1720Leu (VUS); ***PDGFRB***: c.1330C>T p.Pro444Ser (VUS)
SPAST	c.102_122dupC GCCGCCGGGC CGGCCCCTCC	p.Ala35_Pro 41dup	0.0% 0.0%	NA	LP	1	P176	***DCTN1:*** c.3277C>T p.Leu1093Leu (VUS); ***VPS13A:*** c.1661T>C p.Leu554Pro (VUS); ***DNAJC13:*** c.6481A>C p.Lys2161Gln (VUS)
SPG7	c.1807G>A	p.Ala603Thr	0.0% 0.0%	rs370852816	LP	1	P219	***RNASEH2B***: c.529G>A p.Ala177Thr (LP); ***SYNE1***: c.16670A>G p.Asn5557Ser (VUS)
GBA	c.1346C>T	p.Thr410Met	0.0% 0.0%	rs757930613	LP	1	P200	***POLG***: c.3425G>A p.Arg1142Gln (VUS); ***SLC52A3:*** c.1319C>T p.Ser440Leu (VUS)
ADAR	c.2964C>G	p.His988Gln	0.0% 0.0%	NA	LP	1	P202	***ZFYVE26***: c.2663A>G p.Glu888Gly (VUS)
GLB1	c.1174C>G	p.Leu392Val	0.0% 0.0%	NA	LP	1	P198	***GBE1***: c.45G>T p.Glu15Asp (VUS)
GLB1	c.817T>C c.818G>T	p.Trp273Leu	0.0% 0.0%	rs775958270 rs72555362	LP	1	P216	-
HEXA	c.1496G>A	p.Arg499His	0.0% 0.0%	rs121907956	LP	1	P183	***AMPD2***: c.1667G>A p.Arg556His (VUS)

Table 9| List of pathogenic and likely pathogenic variants found in Panel 1. Gene name, nucleotide change, amino acidic change, frequency in the population databases, dbSNP ID, ACMG verdict, the n° of patients and second variants within the patients are shown. **P**: pathogenic; **LP**: likely pathogenic; **VUS**: variants of uncertain significance; **NA**: not available. *MAF (minor allele frequency) in GnomAD and 1000 genome databases

21.4.2 Variants of uncertain significance

Among the variants found in Panel 3, one hundred and fifty-nine (87%) were of uncertain significance according to the ACMG criteria (Figure 25). Twenty-three variants were found in genes causative for ALS (*ERBB4*, *FUS*, *GLE1, HNRNPA1, KIF5A* and *SPG11*) or that confer susceptibility to ALS (*CYP27A2*, *DAO*, *DCTN1*, *EPHA4*, *FGGY*, *NEFH*, *PON2*, *PRPH, SARM1* and *TAF15*), twenty-nine variants were found in genes associated to HSP (*AMPD2*, *AP4M1*, *B4GALNT1*, *CPT1C*,

CYP2U1, *DDHD1*, *ERLIN1*, *FA2H*, *FLRT1*, *HSPD1*, *KIF1A*, *KLC4*, *L1CAM*, *NT5C2*, *RAB3GAP2*, *RTN2*, *SPAST*, *SPG7*, *SPG20, TECPR, ZFR* and *ZFYVE26*) and twenty-seven were found in PD-associated genes (*ADORA1, ATP13A2, DNAJC6, DNAJC13, GIGYF2, PRKN, PDGFRB, PLA2G6, PODXL, POLG, RIC3, TH and VPS13C*). As is Panel 2, also in this case half of the VUS (n = 80) were instead present in genes associated with other NDs, particularly to neuromuscular diseases AD inherited or to diseases with an AR pattern of inheritance that share some clinical features with ALS.

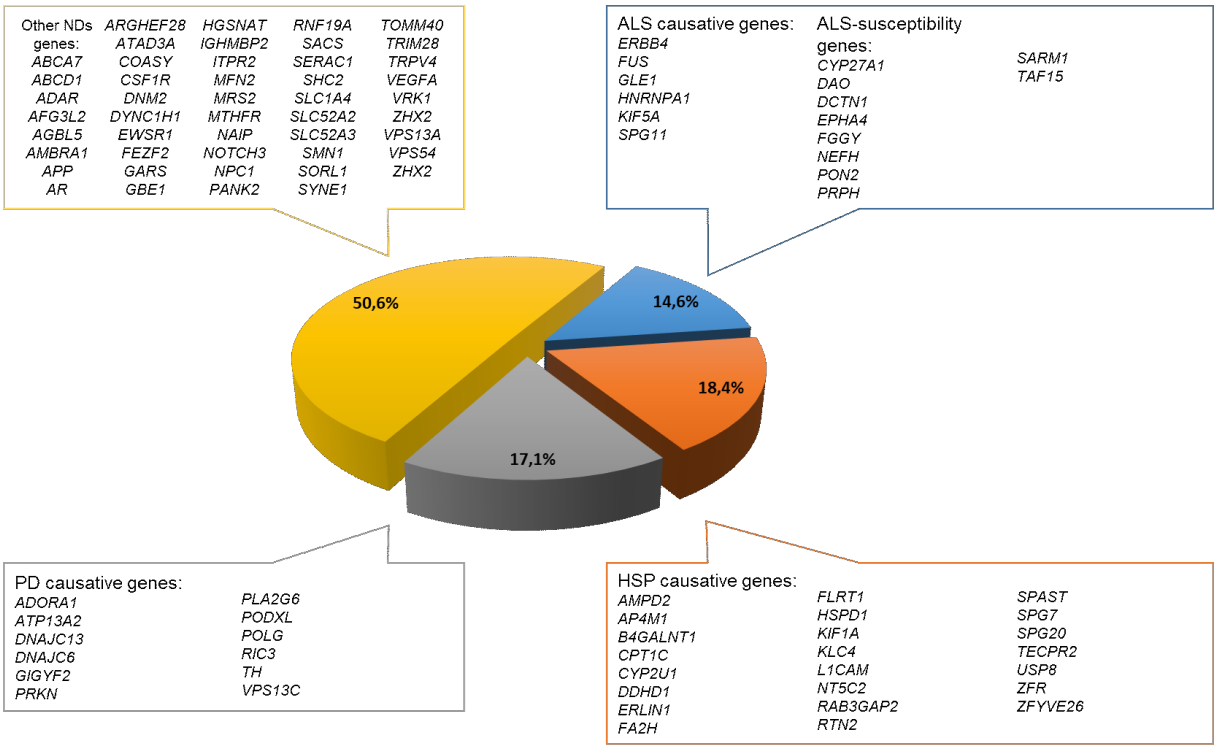

Figure 26| graphic categorization of the VUS variants found in Panel 1 based on the belonging group of disease: ALS (blue), HSP (orange), PD (grey), other NDs (yellow).

From a functional point of view, the majority of variants (n = 128) were missense, followed by variants in the splice-site region (n = 14), small inframe indels (n = 12), variants that alter the 5'-UTR region (n = 4) and variants that disrupt the reading frame (n= 1) (Figure 26).

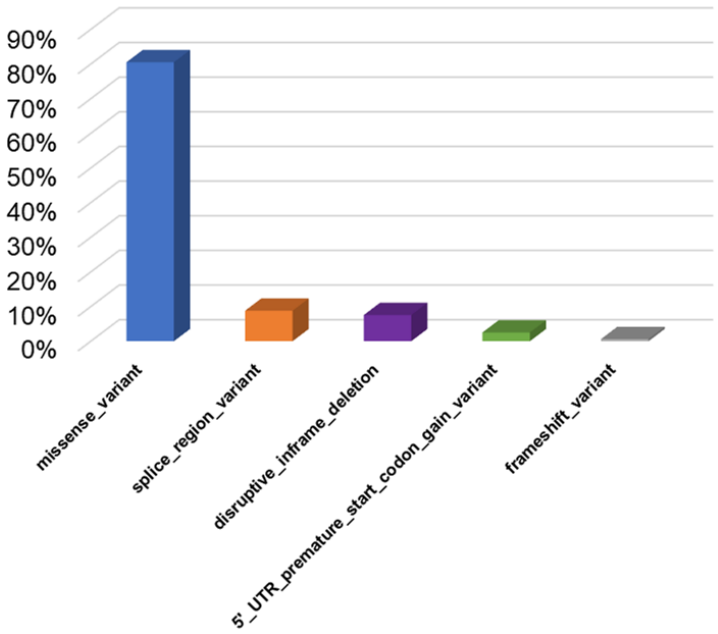

Figure 27| Functional characterization of the VUS variants found in Panel 3: missense variants (blue), splice-region variants (orange), inframe variants (purple), 5'-UTR variants (green) and frameshift variant (grey). X axis represents the relative percentage of the different types of VUS variants found.

21.5 Variant characterization based on pathogenicity and the belonging disease

To better understand the commonalities and the differences among the three gene panels we summarized all the previous results into a unique pie chart per gene panel, in which variants have been subdivided based on i) their pathogenicity and ii) the belonging disease (Figure 27A-C). Pathogenic and likely pathogenic variants constitute the 36% of all the variants in panel 1, while their percentage decreases along panels, representing the 15% and the 12,4% in panel 2 and panel 3 respectively. This decrement is due to the increased number of genes analysed; indeed, the proportion between pathogenic and likely pathogenic ALS variants and the global percentage of pathogenic and likely pathogenic variants does not vary significantly across panels. Additionally, accordingly to patients' phenotype, the percentage of pathogenic and likely pathogenic variants in genes associated with ALS remains the highest across all the three panels, while the percentage of pathogenic and likely pathogenic in genes associated to other NDs increases throughout panels.

On the other hand, looking at VUS, the higher the number of genes analysed the higher the number of VUS, particularly, the number of VUS in genes linked also to other NDs.

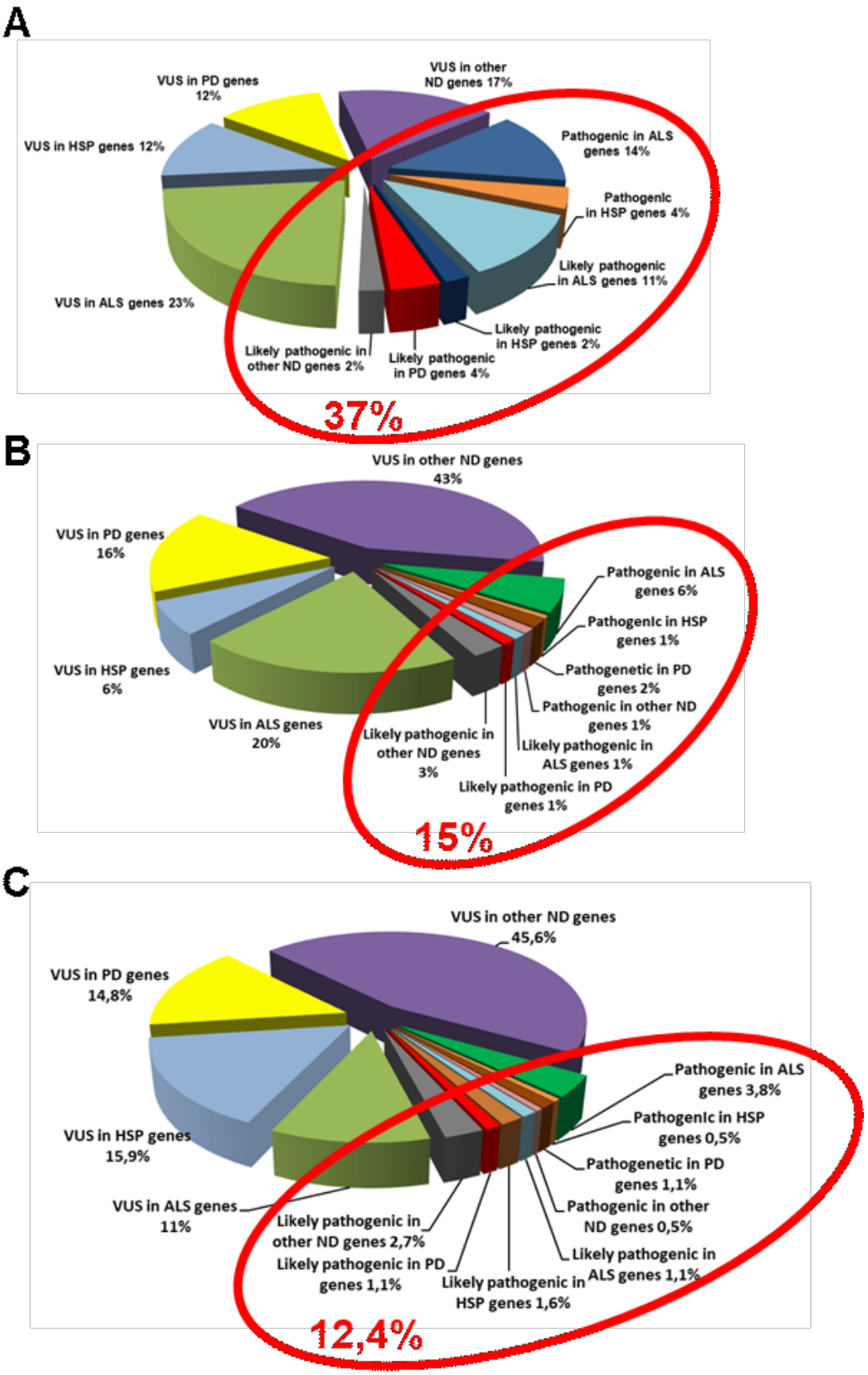

Figure 28| Pie charts show all the variants found subdivided based on the pathogenicity and on the belonging disease. (ALS, HSP, PD and other NDs). A) pie chart of Panel 1; B) pie chart of Panel 2; C) pie chart of Panel 3. Red circle evidences the percentage of the pathogenic and likely pathogenic variants in each gene panel.

21.6 Venn diagrams show pure and "complex" patients' gene profile

We further submitted our cohort of patients to a VENN diagrams webtool (http://bioinformatics.psb.ugent.be/webtools/Venn/) in order to count how many patients harboured one or more than one mutation (either pathogenic, likely pathogenic or VUS) only in ALS-, HSP- and PD genes, or only in genes causative for other NDs (Figure 28A-C). Moreover, Venn diagrams show also "complex patients" such as patients with one or more variants in genes belonging to two or more disease categories.

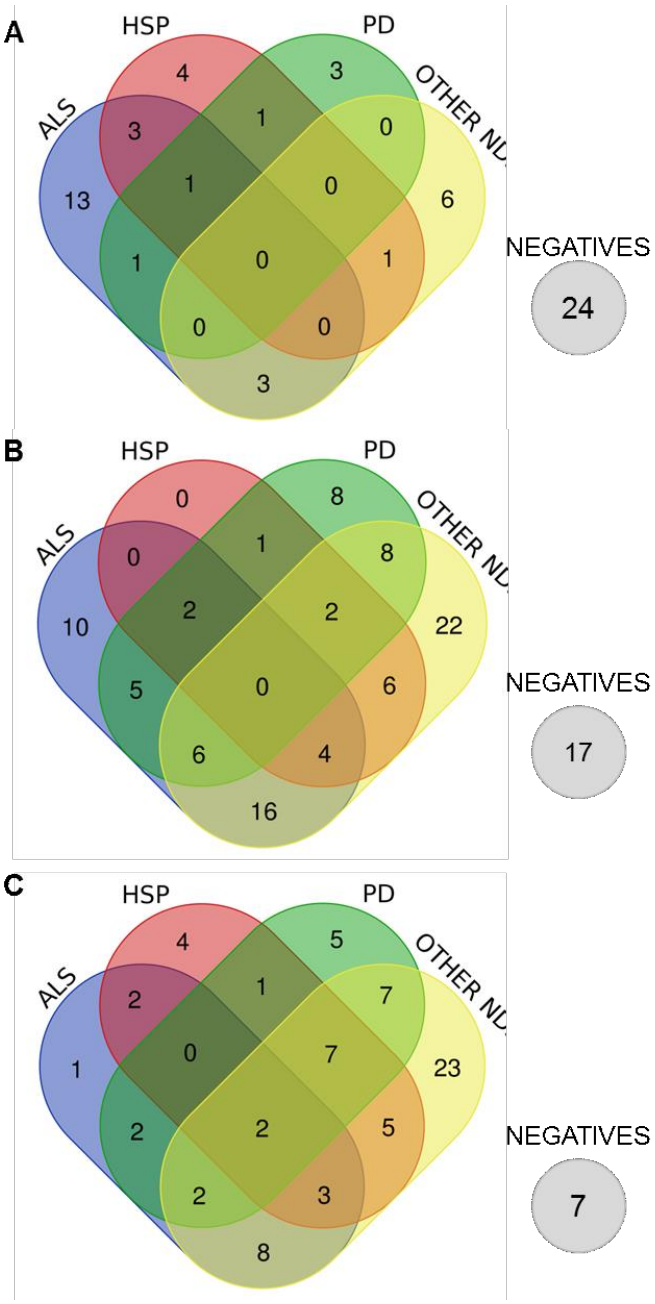

Figure 29| Venn diagrams of panel 1 (A), panel 2 (B) and panel 3 (C) compare the number of patients that have mutations in genes belonging to single diseases (ALS, HSP, PD and Other NDs) and the number of patients that have mutation in more than one disease entity. The grey circle shows the number of patients with no variants.

21.7 Implication of genetic variants in our ALS cohort

Through literature review we generated a list of 726 genetic variants within the 25 genes already described to be implicated in ALS. This list included both single nucleotide changes and small indels with effect on the protein sequence as well as the HRE in the *C9ORF72* gene. Among the 726 variants already implicated in ALS, we found 26 variants within 14 genes across our cohort of 248 SALS cases (Table 10).

			Frequency in our cohort (n=248)		Fisher's exact test (p value)
Gene	**cDNA change**	**Amino acid change**	**N° Heterozygous cases**	**N° Homozygous cases**	**ALS vs GnomAD non-neuro NFE**
ALS2	c.335C>T	p.Ala112Val	1	0	0,0666488
ANG	c.208A>G	p.Ile70Val	1	0	0,431446
C9ORF72	GGGGCC hexanucleotide repeat expansion		16	0	
FIG4	c.122T>C	p.Ile41Thr	2	0	0,266907
FIG4	c.1090A>T	p.Met364Leu	22	0	0,211002
FIG4	c.1961T>C	p.Val654Ala	74	0	1
FUS	**c.521_523+3delGAGGTG**	**p.Gly173_Gly174del**	**3**	**0**	**0,017737**
FUS	c.833-29C>T	splicing	10	0	0,237688
FUS	c.1520G>A	p.Gly507Asp	1	0	
FUS	c.*41G>A	splicing	4	0	0,801982
KIF5A	c.2957C>T	p.Pro986Leu	7	0	0,842912
NEK1	c.782G>A	p.Arg261His	1	0	1
OPTN	c.293T>A	p.Met98Lys	17	0	0,338254
OPTN	c.1401+4A>G	splicing	1	0	0,067209
OPTN	c.1634G>A	p.Arg545Gln	1	0	0,096486
SETX	c.7640T>C	p.Ile2547Thr	4	0	0,360431
SETX	c.4660T>G	p.Cys1554Gly	2	0	0,706525
SETX	c.2975A>G	p.Lys992Arg	9	0	0,863286
SIGMAR1	**c.622C>T**	**p.Arg208Trp**	**4**	**0**	**0,029251**
SOD1	c.260A>G	p.Asn87Ser	1	0	
SOD1	c.272A>C	p.Asp91Ala	1	0	0,254634
SOD1	c.442G>A	p.Gly148Ser	1	0	
SPG11	c.5885A>G	p.Asn1962Ser	1	0	
TARDBP	c.881G>T	p.Gly294Val	1	0	**0,017299**
TARDBP	c.883G>A	p.Gly295Ser	1	0	**0,013002**
VCP	c.572G>A	p.Arg191Gln	1	0	**0,021517**

Table 10| Table show the variants found in our cohort of patients among the 726 variants listed. In bold, the variant that was significantly associated with ALS in our cohort. **NA**: not available; **NFE**, non-Finnish European.

These variants were present among 134/248 (54,4%) cases, all in heterozygous state. Concerning the major ALS genes, this included 16 cases that carried the HRE in the *C9ORF72* gene three missense mutations in *SOD1*, one non-synonymous variant in *VCP*, two missense mutations in *TARDBP* in one case each, and four variants (one missense, two splicing and one in-frame deletion) in the *FUS* gene of 17 patients. Fisher's exact test determined that five variants were significantly over-represented in our cohort of cases compared with GnomAD non-neuro NFE controls:

the p.Arg208Trp variant in the *SIGMAR1* gene (p=0,029251), the variant p.Gly173_Gly174del (p=0,017737), the p.Gly294Val (p=0,017299) and the p.Gly295Ser (p=0,013002) in the *TARDBP* gene, and the p.Arg191Gln in the *VCP* gene (p=0,021517). Among the remaining ALS-implicated variants identified, the p.Ala112Val in *ALS2* and the c.1401+4A>G in *OPTN* had a limited evidence of association but did not reached statistical significance. Since the goal of this association study was to find low-frequency variants associated with our cohort of ALS cases, and the classification of “low frequent” variant relies on the MAF > 5% in the control population,[314] the *FIG4* p.Val654Ala and p.Met364Leu were excluded by further analysis. Eventually, 18 out of 248 (7,25%) of our ALS cases were found to be carrier of two of the reported ALS-implicated variants suggesting oligogenic bases.

22. LATE-ONSET POMPE DISEASE

22.1 LOPD patients

For this study, 90 LOPD patients have been recruited by the Italian centres of the GSDII Group; the age of onset, the Walton-Gardner-Medween score (WGMS), and the disease duration have been recorded (Table 11). The average age of onset of patients was 32,9 years (SD = 16,4), however, patients' age of onset ranged from 3 to 76 years old. The mean WGMS was 3,23 (SD = 2) with a range of 0-7 and the mean disease duration was 12 years (SD = 12,6).

		number	percent (%)
Patients		90	
Gender	Male	39	43
	Female	51	57
Age of onset (years)	Average	32,9	
	SD	16,4	
	Range	3-70	
Disease duration	Average	12	
	SD	12,6	
	Range	2-46	
WGMS	Average	3,2	
	Range	0-7	

Table 11| Epidemiological data of the 90 LOPD cases. The number of cases and the relative percentage have been shown. For the age of onset, the mean age at diagnosis and the standard deviation have been calculated. **WGMS:** Walton-Gardner-Medween score; **SD:** standard deviation.

22.2 GAA genetic screening

The genetic information of 90 patients affected by LOPD has been obtained for the *GAA* gene. As expected, almost all patients (93%, n = 84) harboured the c.-32-13T>G intronic variant, which is the commonest variant among LOPD patients in the Caucasian population, on one allele. The most frequent variants found in compound heterozygosity with the c.-32-13T>G were the c.525delT, present in 15 patients (17%), the c.2237G>A nucleotide change, carried by 10 patients (11%), which introduces a stop-codon and thus it is considered a severe allele, and the c.1465G>A variant, harboured by 9 (10%) patients, which instead is a missense variant that results in the p.Asp489Asn amino acidic change and it is considered a mild mutation that decreases the enzymatic activity of the GAA enzyme.

22.3 MLPA analysis

In 11 out 90 patients (12%), the genetic screening of *GAA* failed to identify the second mutation thus MLPA analysis was performed to look for gross deletions and/or insertions that may represent the second mutation responsible for the pathological phenotypes. Through MLPA analysis we found a heterozygous deletion of exon 17 in one patient (Figure 29A), a heterozygous deletion of exons 9 and 10 in three patients (Figure 29B), and a heterozygous deletion of exon 1 in another patient (Figure 29C) with respect a control DNA (Figure 29D). The heterozygous deletion of exon 1 was also investigated in a family member of our index cases who resulted to be a carrier.

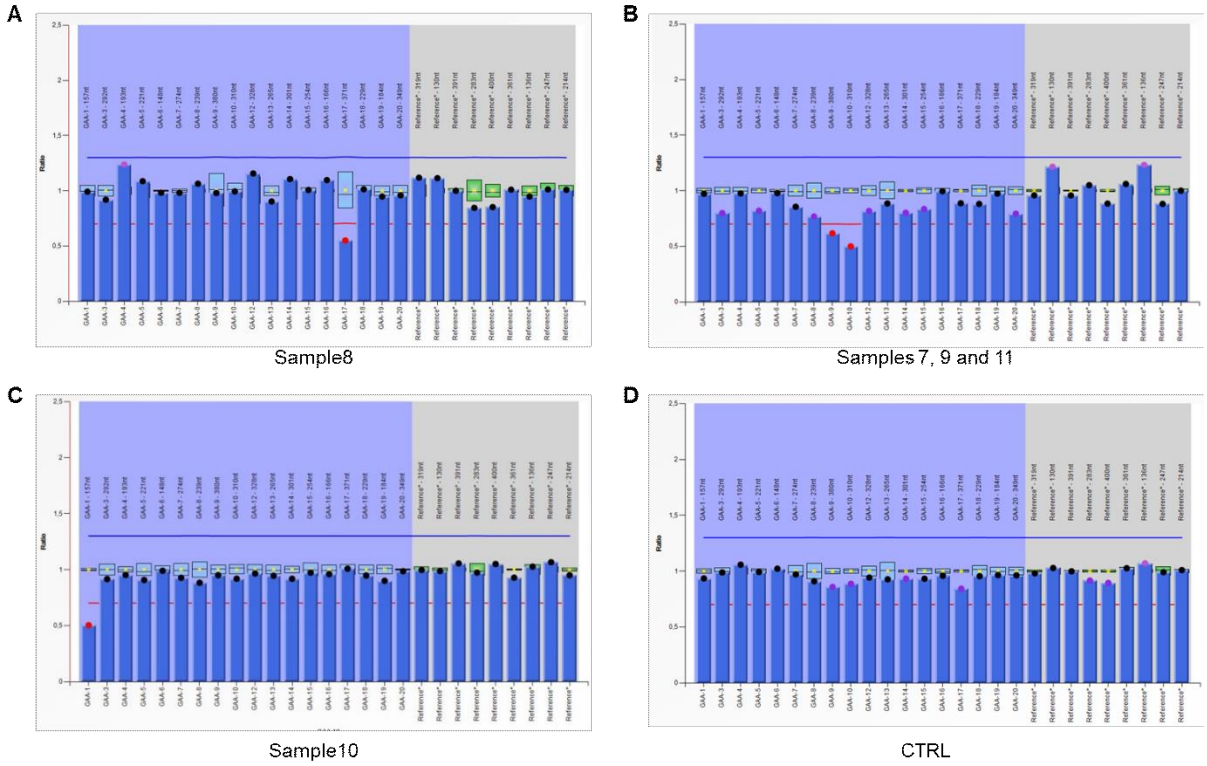

Figure 30| MLPA results indicate a heterozygous deletion of exon 17 in sample 8 (A), a heterozygous deletion of exons 9 and 10 in samples 7, 9 and 11 (B) and a heterozygous deletion of exon 1 in sample 10 (C).

The deletions found, except for the one involving exon 17 due to the lack of the patient's DNA, were validated through Real-time PCR. While the deletion involving exons 9 and 10 resulted to be a false positive (Figure 30A-B), the heterozygous deletion of exon 1 was confirmed in both the index case and in the family member (Figure 30C). All the three DNAs that gave false positive results had a 260/230 absorbance ratio below 1. When this ratio is smaller than 1.8, it indicates contamination probably caused by organic compounds or chaotropic agents, which

absorb at 230 nm. Accordingly, contaminants remaining after DNA extraction, including NaCl or KCl (>40 mM) and other salts, phenol, ethanol, heparin and EDTA (>1.5 mM) may influence MLPA performance. Hence, an additional purification step would have been necessary to correctly analyse these samples.

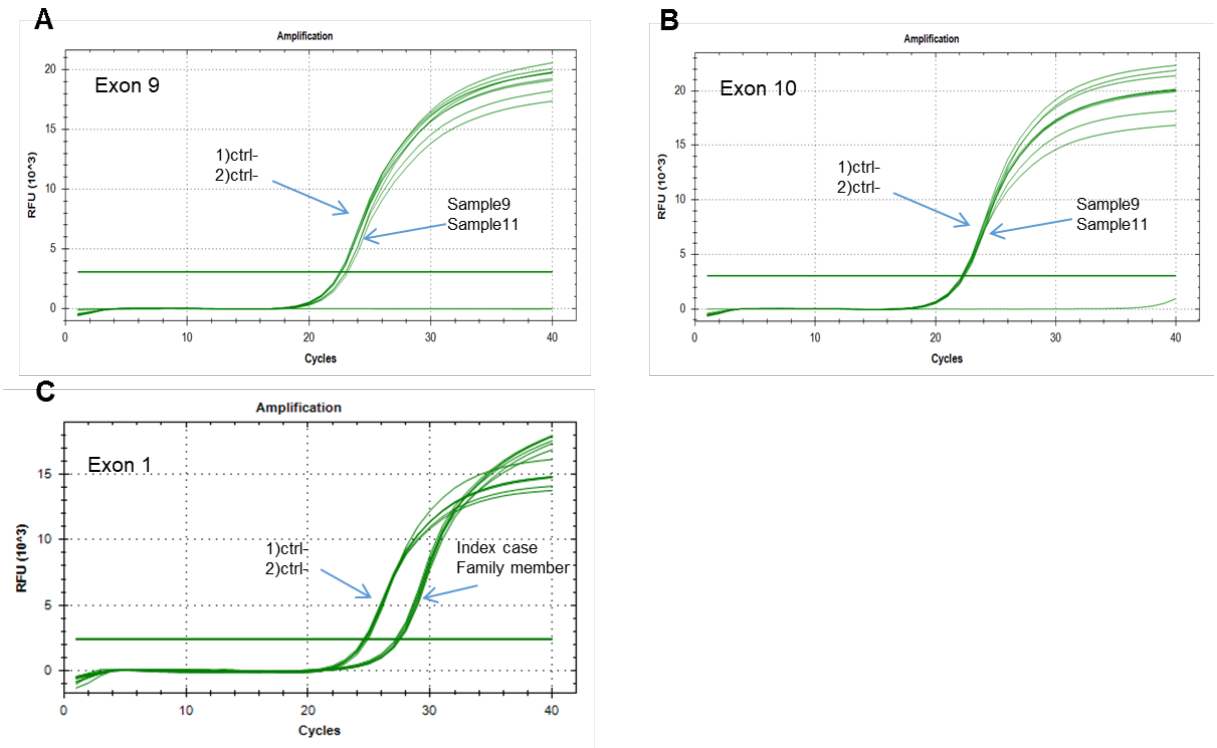

Figure 31| Gene dosage by Real-Time PCR to validate MLPA results. A) gene dosage of exon 9 of the *GAA* gene in cases and controls; B) gene dosage of exon 10 of the *GAA* gene in cases and controls; C) gene dosage of exon 1 of the *GAA* gene in cases and controls.

22.4 GAA SNP analysis

To date, within the GAA gene have been reported up to 765 mutations (http://cluster15.erasmusmc.nl/klgn/pompe/mutations.html, last update October 12, 2020), and 16 single nucleotide polymorphisms (SNPs) have been reported within the coding region.[315] Among the latters, some alter the amino acidic sequence, and thus they may have some effect on the enzyme, for instance on protein stability and 3D-structure. Hence, we selected three SNPs within the *GAA* gene: i) c.596A>G (p.His199Arg; rs1042393) in exon 3; ii) c.668A>G (p.Arg223His; rs1042395) in exon 3; iii) c.2338A>G (p.Val780Ile; rs1126690) in exon 18, and we investigated their presence in our cohort of LOPD patients. The purpose of this analysis was to see whether these SNPs may correlate with the clinical variability of patients, in association with their *GAA* genotypes.

Firstly, through 3D-modeling, we obtained the 3D model of the GAA enzyme and the relative position of the SNPs understudied (Figure 31), obtained using models are available from the PDB database.

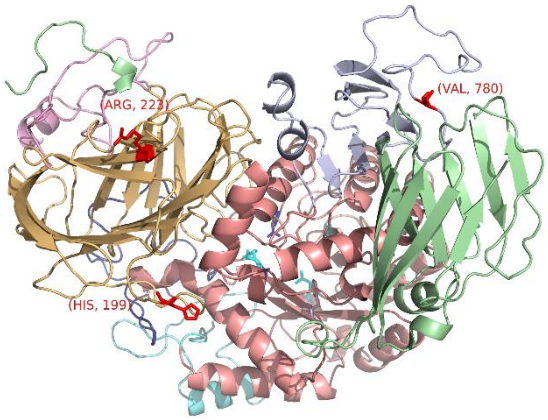

Figure 32| 3D model of the GAA enzyme and the relative position of the SNPs

All the three SNPs are located at the protein surface and far from the catalitic site of the enzyme. Through DUET, mCSM and SDM tools, the protein stability free energy change (DDG) upon single SNP insertion has been obtained to investigate the effect on the protein stability (Table 12). SNPs were considered destabilizing if DDG<-0.5 kilocalorie per mole (Kcal/mol), stabilizing if DDG>0.5 Kcal/mol and neutral if -0.5<=DDG<=0.5 Kcal/mol.[316]

	mCSM	DUET	SDM
p.His199Arg	0.203 Kcal/mol (Neutral)	0.407 Kcal/mol (Neutral)	-0.04 Kcal/mol (Neutral)
p.Arg223His	-1.817 Kcal/mol (Destabilizing)	-2.045 Kcal/mol (Destabilizing)	-0.47 Kcal/mol (Neutral)
p.Val780Ile	-0.627 Kcal/mol (Destabilizing)	-0.574 Kcal/mol (Destabilizing)	-0.03 Kcal/mol (Neutral)

Table 12| The table shows the Kcal/mol of the three SNPs according to DUET, mCSM and SDM tools.

The p.Arg223His and the p.Val780Ile resulted destabilizing according to mCSM and DUET tools, while the p.His199Arg were considered neutral according to all the three tools queried. According to InterPro, the three SNPs studied were not involved in potential interaction surfaces, and no changes in secondary structure due to the SNPs, either alone or in combination, has been observed by using DSST tool.

Subsequently, we assessed the frequencies of the three SNPs in our cohort of patients:

1) The c.596A>G nucleotide substitution was found in 64 patients (73,6%). This variant has been observed in the 74% of the European non-Finnish population (GnomAD)
2) The SNP c.668A>G was found also in 64 patients (73,6%) as the previous one and showed the same allele frequency also in the European non-Finnish population. Since the two SNPs reside within the same exon they may belong to the same haplotype and segregate together
3) The c.2338A>G nucleotide substitution was found in 68 patients (78,2%), in line with the frequency observed in the European non-Finnish population (GnomAD) which is 75,6%.

Unfortunately, no correlations with the clinical phenotype of patients have been found in association with the SNPs analysed.

22.5 Exome for genetic variability

22.5.1 Clinical cases and variant filtering

Among our cohort of LOPD patients, we selected 30 subjects genotypically homogeneous for performing WES analysis with the aim to identify variants that may have an impact on the clinical variability in LOPD. To reduce genotype variability, we included in the subset of patients only individuals with the c.-32-13T>G on one allele, and the following mutation on the second GAA allele: c.525delT, c.2237G>A, c.2481+102_2646+31del, c.1551+1G>C, c.1802C>G, c.2219_2220delTG, c.784G>A, c.1927G>A, c.1465G>A and c.1655T>C. The choice has been made on the bases of the frequency of these mutations in the Italian population and on the presence of these variants in our cohort of patients. In table 13, the *GAA* genotype and the principal clinical features for each patient are shown.

	Gender	Age at symptom onset (years)	Age at diagnosis (years)	*GAA* genotype	WGMS	6MWT (meters)	FVC%	Ambulation support	Respiratory Support	ERT (year)	ERT responder
BR01	F	39	41	c.-32-13T>G/c.2237G>A	3	424	105	no	no	yes (2011)	yes
FB01	M	44	45	c.-32-13T>G/c.2237G>A	6	--	42	yes	yes	yes (2011)	yes
FR01	M	26	28	c.-32-13T>G/c.2237G>A	1	375	86	no	no	yes (2005)	yes
GS01	F	15	15	c.-32-13T>G/c.784G>A	5	216	67	no	no	yes (2006)	yes
BA51	M	49	63	c.-32-13T>G/c.1927G>A	3	125	35	no	yes	yes (2014)	yes
DD45	M	30	66	c.-32-13T>G/c.1927G>A	3	303	54	no	yes	yes (2012)	no
MN01	M	43	45	c.-32-13T>G/c.1551+1G>C	2	486	85	no	yes	yes (20100)	yes
MN02	F	38	54	c.-32-13T>G/c.1551+1G>C	2	451	55	no	yes	yes (2011)	yes
VG01	M	10	31	c.-32-13T>G/c.1927G>A	3	375	39	--	--	yes (NA)	yes
IM01	M	65	71	c.-32-13T>G/c.525delT	2	527	86	yes	yes	yes (2013)	yes
DA01	M	27	29	c.-32-13T>G/c.1121G>T	2	640	77	no	no	yes (2009)	yes
DR01	F	20	28	c.-32-13T>G/c.1802C>G	3	284	53	yes	yes	yes (2008)	yes
SA01	F	35	50	c.-32-13T>G/c.784G>A	3	275	72	yes	no	yes (2015)	yes
ZA01	F	15	15	c.-32-13T>G/c.1927G>A	0	665	91	no	no	yes (2011)	yes
US01	F	40	51	c.-32-13T>G/c.525delT	3	390	97	no	yes	yes (2012)	yes
GG01	F	26	37	c.-32-13T>G/c.784G>A	1	563	1	no	no	yes (2013)	yes
PT01	F	--	43	c.-32-13T>G/c.1465G>A	3	249	45	yes	yes	yes (2007)	yes
ME01	F	--	35	c.-32-13T>G/c.1465G>A	6	--	82	yes	no	yes (2010)	yes
SM01	F	45	76	c.-32-13T>G/c.784G>A	3	50	47	no	yes	yes (2016)	yes
TL55	M	25	57	c.-32-13T>G/c.1655T>C	--	--	65	no	no	yes (2012)	yes
FM63	F	--	36	c.-32-13T>G/c.2237G>A	--	114	55	yes	yes	yes (2006)	no
CA01	M	35	46	c.-32-13T>G/c.2237G>A	7/10	300	43	no	yes	yes (2007)	yes
MN03	F	25	52	c.-32-13T>G/c.1551+1G>C	2	320	46	no	yes	yes (2012)	yes
RA01	F	38	39	c.-32-13T>G/c.2481+102_2646+31del	5	10	81	yes	no	yes (2015)	--
GAA81	F	13	13	c.-32-13T>G/c.2237G>A	8	150	53	yes	yes	yes (2006)	no
SM02	M	--	--	c.-32-13T>G/c.2237G>A	--	--	--	no	no	no	--
SM04	M	--	--	c.-32-13T>G/c.2237G>A	--	--	--	no	no	no	--
BR04	M	--	--	c.-32-13T>G/c.2237G>A	--	--	--	no	no	no	--
BGA01_II	F	25	23	c.-32-13T>G/c.525delT	3	330	94	no	no	yes (2010)	yes
BGA02	F	8	37	c.-32-13T>G/c.525delT	5-6	284	77	yes	no	yes (2006)	yes

Table 13| for each patient, genotype, age at symptom onset, age at diagnosis, *GAA* genotype, WGMS (Walton-Gardner-Medween score), 6 MWT (minute walking test), percentage of FVC (forced vital capacity), ambulation and respiratory support, enzymatic replacement therapy (ERT) and response to therapy are shown. **NA**: not available.

After WES analysis, hundreds of thousands of variants for each patient have been collected, thus the first step of variant filtering has been performed. Specifically, were retained variants with i) a functional effect and thus that alter the amino acidic sequence, ii) MAF < 0,05 in the NFE population (GnomAD, ExAC). We choose a MAF of 0,05 because in this part of the project the aim was to find variants associated with LOPD, including rare variants and not frequent polymorphisms, while in the ALS study we looked for possible pathogenic variants, that usually are rarer compared to polymorphisms, and thus with a MAF < 0,01; iii) variants with a read depth ≥ 10 reads (Table 14). Upon variants' filtering, about 1000-1500 variants per sample have been retained for a total of 37904 variants.

Sample	Gender	Total variants (n)	Variants within coding sequence (n)	Removal Synonymous variants (n)	MAF < 0,05 gnomAD NFE	MAF < 0,05 NFE ExAC	Read Depth ≥10
BA51	M	149548	25408	13337	1664	1341	1329
BGA01_II	F	186427	25170	13180	1704	1339	1334
BGA02	F	157235	25106	13170	1724	1387	1378
BR01	F	180426	26196	13975	1943	1571	1563
BR04	M	149899	26136	13907	1884	1538	1530
CA01	M	196634	25462	13279	1642	1308	1300
DA01	M	281057	26047	13695	1871	1481	1466
DD45	M	169623	25838	13663	1727	1364	1355
DR01	F	253066	25956	13747	1966	1575	1569
FB01	M	299373	25270	13322	1722	1357	1348
FM63	F	295725	26201	13790	1782	1441	1429
FR01	M	319336	25343	13337	1747	1385	1364
GAA81	F	294141	26530	14052	1942	1545	1539
GG01	F	247706	26048	13749	1958	1564	1553
GS01	F	228670	26162	13813	1393	1112	1103
IM01	M	324621	26873	13721	1321	1035	1030
ME01	F	153291	25594	13428	1313	1046	1042
MN01	M	165292	25888	13639	1489	1235	1227
MN02	F	173641	25854	13565	1433	1133	1128
MN03	F	160812	26163	13761	1446	1185	1178
PT01	F	157410	25385	13330	1336	1075	1070
RA01	F	231920	26164	13863	1377	1074	1068
SA01	F	283844	26201	13866	1416	1119	1112
SM01	F	253364	26156	13908	1444	1152	1146
SM02	M	161295	25719	13664	1431	1086	1081
SM04	M	186131	26103	13861	1390	1069	1063
TL55	M	254536	25994	13702	1268	1035	1022
US01	F	257324	25972	13731	1294	1011	1005
VG01	M	308566	26000	13707	1344	1068	1063
ZA01	F	270598	25541	13476	1869	1521	1509
Total variants		**6751511**	**776480**	**409238**	**47840**	**38152**	**37904**

Table 14| Variant filtering workflow for each patient is shown. **NFE**: non-Finnish Europeans.

22.5.2 Pathway and gene selection

We decided to focus our analysis only on genes belonging to pathways involved in the pathogenesis of LOPD, selected using KEGG (https://www.genome.jp/kegg/), WikiPathways (https://www.wikipathways.org/index.php/WikiPathways) and Reactome (https://reactome.org/). Thus, we focused on: ACE inhibitor pathway,[317] apoptosis,[318,319] Regulation of autophagy,[274] Glycogen metabolism,[320] Lysosome[274] and Renin-angiotensin system.[317] In total 302 genes have been analysed: 113 genes for the Lysosome pathway, 39 for the Regulation of autophagy pathway, 85 for the Apoptosis pathway, 21 for the Renin-angiotensin system pathway, 36 for the Glycogen metabolism pathway and 8 for the ACE inhibitor pathway (Figure 32).

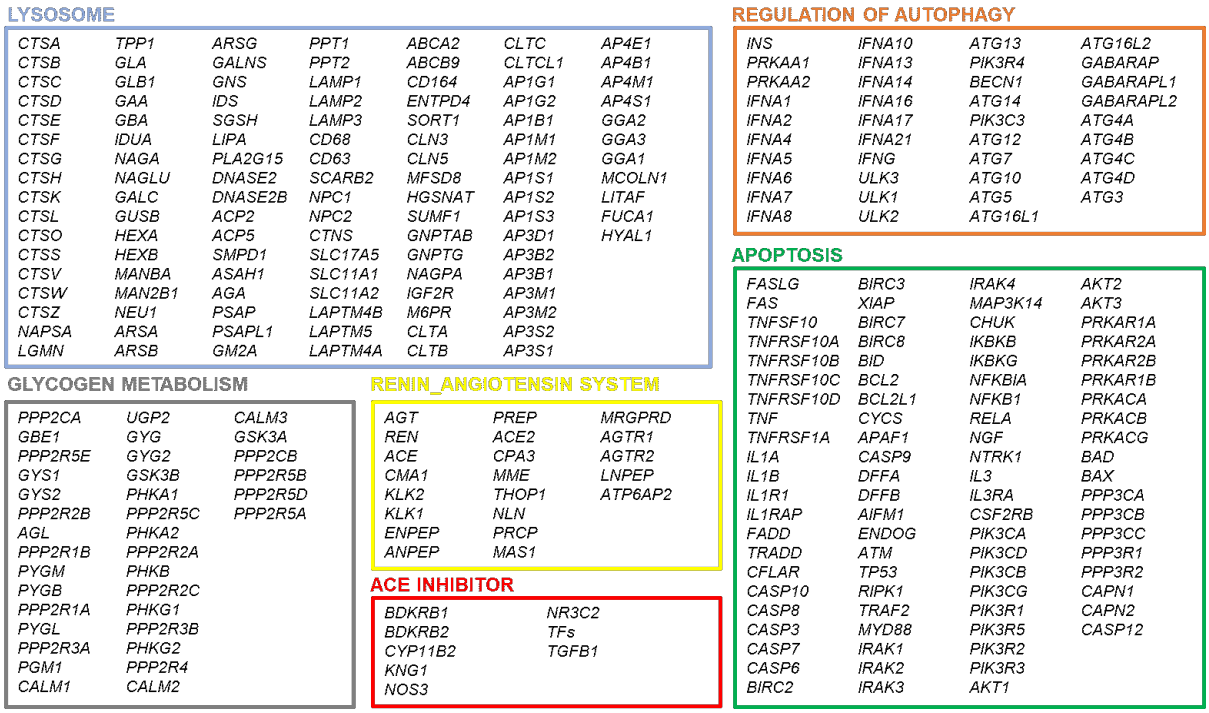

Figure 33| graphic representation of the genes analysed in this study. Genes involved in the Lysosome pathway are in blue, genes involved in Regulation of autophagy pathway are in orange, genes involved in Glycogen metabolism pathway are in grey, genes involved in Renin-angiotensin system pathway are in yellow, genes involved in the ACE inhibitor pathway are in red and genes involved in the Apoptosis pathway are in green.

We found a total of 176 variants in 103 (34%) out of 302 genes analysed. Specifically, we found 82 variants in 46 genes of the Lysosome pathway, 31 variants in 15 genes linked to the Regulation of autophagy pathway, 45 variants in 28 genes related to the Apoptosis pathway, 5 variants in 5 genes of the Renin-angiotensin pathway, 7 variants in 5 genes of the Glycogen metabolism pathway and 6 variants within 4 genes of the ACE inhibitor pathway (Figure 33).

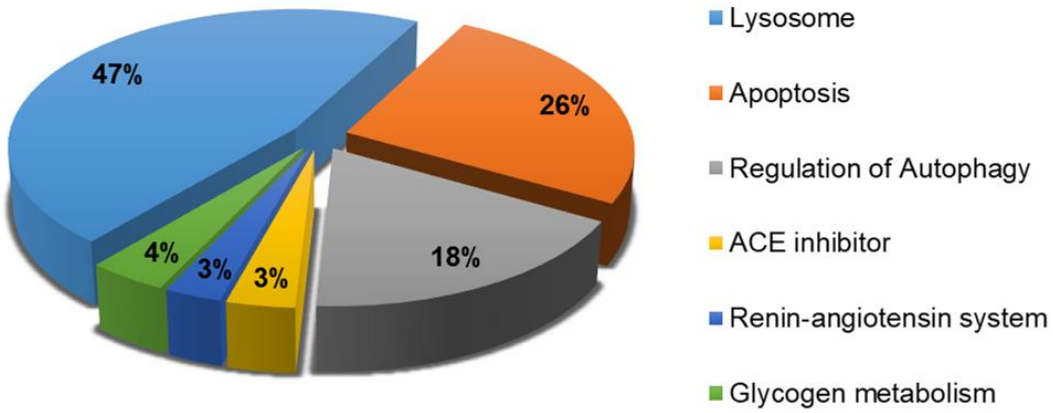

Figure 34| Pie chart representing the variants' distribution across the six pathways analysed.

22.5.3 Associated-LOPD variants in genes of the Lysosome pathway

To test whether the functional variants found within the genes linked to the Lysosome pathway were associated with LOPD, Fisher's exact test has been performed, comparing the frequency of each variant to the frequency of the non-Finnish European population in GnomAD (Table 15). Among the 82 variants found in the 113 genes connected to this pathway, 36 resulted significantly associated with LOPD, comprising the pathogenic variants in the *GAA* gene. The significantly associated variants have been found in genes encoding for the adaptor related protein complex subunits that are involved in the formation of clathrin-coated synaptic vesicles, as well as in the *CLTCL1* that codifies for the clathrin heavy chain-like 1. Three variants were found in the *FUCA1* gene which codifies for a lysosomal enzyme involved in the degradation of fucose-containing glycoproteins and glycolipids and it is connected to Fucosidosis which is a lysosomal storage disorder, and other two variants were in the *NAGA* gene which is also involved in another lysosomal storage disorder.

			Frequency in our cohort (n=30)		Fisher's exact test (p value)
Gene	**cDNA change**	**Amino acid change**	**N° Heterozygous cases**	**N° Homozygous cases**	**LOPD vs GnomAD NFE**
AGA	c.313C>A	p.Leu105Ile	1	0	0,632702
AGA	c.34G>T	p.Val12Leu	1	0	0,984808
AP1M2	c.254A>G	p.Tyr85Cys	1	0	0,35695
AP1S3	c.11T>G	p.Phe4Cys	1	0	0,657474
AP1S3	c.95C>T	p.Thr32Ile	1	0	0,320118
AP3B1	c.2661C>A	p.Phe887Leu	2	0	0,093588
AP3B1	c.2409_2411del	p.803_804del	1	0	NA
AP3B1	c.2995G>A	p.Val999Met	1	0	0,355766
AP3B2	c.2543C>T	p.Pro848Leu	1	0	**0,00095**
AP3B2	c.2396G>C	p.Ser799Thr	1	0	**0,002122**
AP3M2	c.104C>G	p.Ala35Gly	1	0	**0,028154**
AP3S1	c.70dupA	p.Tyr23fs	1	0	NA
AP4B1	c.1497G>T	p.Leu499Phe	1	0	**0,004216**
AP4B1	c.767C>T	p.Thr256Ile	1	0	0,081168
AP4E1	c.2040T>G	p.Ile680Met	1	0	**0,003258**
AP4E1	c.415G>A	p.Val139Ile	1	0	**0,001862**
ARSA	c.20G>T	p.Arg7Leu	1	0	**0,00322**
ARSG	c.1192C>T	p.Arg398Trp	1	0	1
ARSG	c.1154G>A	p.Arg385His	2	0	0,47748
ASAH1	c.372T>A	p.Asp124Glu	1	0	1
CD68	c.1018G>A	p.Ala340Thr	1	0	1
CLN3	c.1211A>G	p.His404Arg	2	0	0,547676
CLN5	c.61C>T	p.Pro21Ser	1	0	0,337704
CLN5	c.1166A>C	p.Lys389Thr	1	0	**0,008404**
CLTCL1	c.1945G>A	p.Glu649Lys	2	0	**0,000002**
CLTCL1	c.3058G>A	p.Glu1020Lys	2	0	**0,000002**
CLTCL1	c.130G>T	p.Val44Phe	1	0	0,416654
CLTCL1	c.3068A>T	p.Asn1023Ile	1	0	1
CLTCL1	c.3097A>T	p.Ile1033Phe	2	0	**0,000002**
CLTCL1	c.4603G>C	p.Val1535Leu	2	0	NA
CLTCL1	c.4721C>T	p.Pro1574Leu	2	0	**0,000002**
CLTCL1	c.5C>T	p.Ala2Val	1	0	0,092052
CLTCL1	c.2239A>T	p.Ile747Leu	1	0	0,207404
CLTCL1	c.182C>T	p.Pro61Leu	2	0	**0,03555**
CLTCL1	c.4603G>A	p.Val1535Met	2	0	NA
CLTCL1	c.4688G>A	p.Arg1563His	2	0	NA
CLTCL1	c.614A>G	p.Lys205Arg	2	0	**0,03598**
CTSC	c.1357A>G	p.Ile453Val	3	0	0,250258
CTSD	c.844G>A	p.Gly282Arg	1	0	0,37985
CTSE	c.244A>G	p.Ile82Val	1	0	0,884988
CTSH	c.479A>G	p.Lys160Arg	2	0	0,375882
CTSO	c.359T>C	p.Val120Ala	1	0	**0,002326**
CTSW	c.317G>A	p.Arg106Gln	1	0	0,069874
FUCA1	c.34G>T	p.Gly12Cys	1	0	**0,007412**
FUCA1	c.37C>T	p.Pro13Ser	1	0	**0,007562**
FUCA1	c.7G>C	p.Ala3Pro	2	0	**0,045674**
GAA	c.1927G>A	p.Gly643Arg	4	0	**<0,000001**
GAA	c.2237G>A	p.Trp746Ter	5	0	**<0,000001**
GAA	c.525delT	p.Thr175fs	4	0	NA
GAA	c.784G>A	p.Glu262Lys	3	0	**<0,000001**
GAA	c.1124G>T	p.Arg375Leu	1	0	**0,001608**
GBA	c.1223C>T	p.Thr408Met	1	0	0,496474
GGA1	c.1213G>A	p.Glu405Lys	1	0	0,828408
GLB1	c.325C>T	p.Arg109Trp	2	0	0,286882
GLB1	c.1306C>T	p.Leu436Phe	1	0	0,734594
GM2A	c.55G>A	p.Ala19Thr	1	0	1
GNPTAE	c.3598G>A	p.Glu1200Lys	1	0	**0,042296**
GNPTG	c.67G>C	p.Gly23Arg	1	0	**0,024354**
GUSB	c.1504A>G	p.Ser502Gly	1	0	NA
LAMP1	c.611C>T	p.Ala204Val	2	0	**0,000464**
LAMP1	c.644A>G	p.Lys215Arg	1	0	**0,001404**
LGMN	c.995C>T	p.Thr332Met	1	0	0,725102
IGF2R	c.2156G>C	p.Arg719Thr	2	0	**0,000222**
IGF2R	c.6059A>G	p.Asn2020Ser	2	0	1,310682
M6PR	c.218_223del	p.73_75del	1	0	NA
MANBA	c.2482G>A	p.Val828Ile	1	0	0,820838
MCOLN1	c.305G>A	p.Arg102Gln	1	0	**0,039096**
MCOLN1	c.782C>T	p.Thr261Met	2	0	**0,01152**
MCOLN1	c.856C>A	p.His286Asn	1	0	**0,017132**
MFSD8	c.1268C>T	p.Ala423Val	1	0	0,686836
NAGA	c.697G>A	p.Val233Met	2	0	**0,000108**
NAGA	c.973G>A	p.Glu325Lys	2	0	**0,012918**
NAGPA	c.455T>C	p.Val152Ala	1	0	**0,029586**
NPC1	c.3797G>A	p.Arg1266Gln	2	0	0,466692
NPC2	c.88G>A	p.Val30Met	1	0	0,06332
NPC2	c.212A>G	p.Lys71Arg	1	0	**0,032246**
PLA2G1	c.887G>C	p.Arg296Pro	1	0	**0,01113**
SLC11A	c.1328T>C	p.Val443Ala	1	0	0,55399
SMPD1	c.8G>A	p.Arg3His	1	0	0,064548
SMPD1	c.1071C>T	p.Ala357Ala	1	0	0,432006
SORT1	c.1330G>C	p.Glu444Gln	2	0	0,202316
TPP1	c.638C>T	p.Ser213Leu	1	0	**0,004182**

Table 15| Table shows the variants found in our cohort of patients among the 113 genes of the Lysosome pathway. In bold, variants significantly associated with LOPD in our cohort. **NA**: not available; **NFE**, non-Finnish European.

22.5.4 Associated-LOPD variants in genes of the ACE inhibition pathway

Six variants were found in 5 out of 8 genes of the ACE inhibition pathway (Table 16). all variants were found in the heterozygous state and only the variant in the *NR3C2* gene reached statistical significance. This gene codifies for a mineralocorticoid receptor, which mediates aldosterone actions on salt and water balance within cells.

			Frequency in our cohort (n=30)		Fisher's exact test (p value)
Gene	**cDNA change**	**Amino acid change**	**N° Heterozygous cases**	**N° Homozygous cases**	**LOPD vs GnomAD NFE**
BDKRB1	c.844C>T	p.Arg282Ter	1	0	0,181888
CYP11B2	c.1016T>C	p.Ile339Thr	2	0	0,068964
CYP11B2	c.1303G>A	p.Gly435Ser	1	0	0,920272
KNG1	c.1290C>G	p.Asp430Glu	1	0	1,187306
NR3C2	c.1705A>G	p.Arg569Gly	1	0	**0,018958**
TF	c.2012G>A	p.Gly671Glu	1	0	0,241594

Table 16| Table shows the variants found in our cohort of patients among the 8 genes of the ACE inhibition pathway. In bold, variants significantly associated with LOPD in our cohort. **NFE**, non-Finnish European.

22.5.5 Associated-LOPD variants in genes of the Glycogen metabolism pathway

Seven variants were found in 5 of the 36 genes involved in the Glycogen metabolism pathway (Table 17).

			Frequency in our cohort (n=30)		Fisher's exact test (p value)
Gene	**cDNA change**	**Amino acid change**	**N° Heterozygous cases**	**N° Homozygous cases**	**LOPD vs GnomAD NFE**
GSK3B	c.930T>A	p.His310Gln	1	0	**0,014454**
GYS2	c.1965G>C	p.Gln655His	4	0	0,098873
PHKB	c.2309A>G	p.Tyr770Cys	8	0	**0,002941**
PPP2R1B	c.1825G>C	p.Val609Leu	1	0	1
PPP2R1B	c.269G>A	p.Gly90Asp	1	0	0,506634
PPP2R1B	c.1867A>T	p.Asn623Tyr	5	0	**0,000272**
PYGL	c.1900G>C	p.Asp634His	1	0	0,322001

Table 17| Table shows the variants found in our cohort of patients among the 36 genes of the Glycogen metabolism pathway. In bold, variants significantly associated with LOPD in our cohort. **NFE**, non-Finnish European.

Among the seven variants, three were significantly associated with LOPD. Specifically, one variant has been found in the *GSK3B* gene which codifies for a glycogen synthase kinase and in skeletal muscle, contributes to insulin regulation of glycogen synthesis by phosphorylating and inhibiting *GYS1* activity, and hence glycogen synthesis. This variant has been found in an individual of our cohort which does not display any symptom yet despite the biallelic mutation in the *GAA* gene, suggesting that this variant may be protective. The other associated variants were found in the *PHKB* gene of eight patients, which is responsible for the glycogen

storage disease type 9B, also known as phosphorylase kinase deficiency of liver and muscle, and in the *PPP2R1B* gene of five patients, which encodes for a regulatory subunit of phosphatase

22.5.6 Associated-LOPD variants in genes of the Renin-angiotensin system pathway

Among the 21 genes involved in the Renin-angiotensin system pathway, we found 5 variants in 5 different genes pathway (Table 18). Three variants resulted significantly associated with our LOPD cohort and fallen in the *AGTR2*, *ENPEP,* and *KLK1* genes. *AGTR2* codifies for the angiotensin II receptor type 2, *ENPEP* codifies for the glutamyl aminopeptidase while the *KLK1* gene encodes for a Kallikrein, which is a serine protease. This variant in the KLK1 gene has been seen in eight different patients (BR01, BR04, CA01, DA01, DR01, IM01, PT01, SM02, SM04), however, the clinical features recorded for these patients are very diverse, suggesting that this variant is not involved in the clinical variability of patients.

			Frequency in our cohort (n=30)		Fisher's exact test (p value)
Gene	**cDNA change**	**Amino acid change**	**N° Heterozygous cases**	**N° Homozygous cases**	**LOPD vs GnomAD NFE**
ACE2	c.2158A>G	p.Asn720Asp	0	1	1
AGTR2	c.231G>T	p.Lys77Asn	0	1	**0,024576**
CMA1	c.197A>G	p.His66Arg	1	0	1
ENPEP	c.817C>T	p.Arg273Ter	1	0	**0,002122**
KLK1	c.230G>A	p.Arg77His	8	0	**0,001399**

Table181| Table shows the variants found in our cohort of patients among the 21 genes of the Renin-angiotensin system pathway. In bold, variants significantly associated with LOPD in our cohort. **NFE**, non-Finnish European.

22.5.7 Associated-LOPD variants in genes of the Regulation of autophagy pathway

31 variants were found in 15 out of 39 genes of the pathway of the Regulation of autophagy (Table 19). Half of the variants (n = 15) were significantly associated with our LOPD cohort. The genes carrying the variants can be classified in genes of the interferon-alpha family (*IFNA10*, *IFNA14*, *IFNA16*, *IFNA17*, *IFNA21*, *IFNA5*, *IFNA6* and *IFNA7*), in genes autophagy-related (*ATG12*, *ATG13*, *ATG4B* and *ATG4C*) and in genes that codify for Unc-51 like autophagy activating kinase (*ULK1* and *ULK2*).

			Frequency in our cohort (n=30)		Fisher's exact test (p value)
Gene	**cDNA change**	**Amino acid change**	**N° Heterozygous cases**	**N° Homozygous cases**	**LOPD vs GnomAD NFE**
ATG12	c.43A>G	p.Ile15Val	1	0	1
ATG13	c.1189G>A	p.Gly397Arg	1	0	NA
ATG4B	c.566G>A	p.Cys189Tyr	1	0	**0,002132**
ATG4B	c.989A>G	p.Asn330Ser	1	0	0,521593
ATG4C	c.976A>C	p.Lys326Gln	1	0	0,121087
ATG4C	c.955_956del	p.Cys319fs	1	0	NA
BECN1	c.658G>A	p.Glu220Lys	2	0	**0,032905**
IFNA10	c.178C>G	p.Arg60Gly	8	0	**0,000081**
IFNA10	c.181A>T	p.Ile61Phe	8	0	**0,000032**
IFNA10	c.354C>A	p.Asp118Glu	1	0	0,523478
IFNA10	c.492G>C	p.Trp164Cys	2	0	0,169623
IFNA10	c.496G>A	p.Val166Ile	1	0	0,330181
IFNA14	c.252G>C	p.Met84Ile	1	0	**0,022968**
IFNA14	c.488C>T	p.Ala163Val	1	0	0,411159
IFNA16	c.139A>G	p.Ile47Val	8	0	**0,004992**
IFNA16	c.146A>C	p.His49Pro	1	0	0,338913
IFNA16	c.194T>A	p.Val65Glu	8	0	**0,005271**
IFNA16	c.283G>C:	p.Asp95His	8	0	**0,005231**
IFNA16	c.371C>T	p.Thr124Ile	8	0	**0,005131**
IFNA16	c.395T>C	p.Ile132Thr	8	0	**0,004239**
IFNA16	c.397G>C	p.Ala133Pro	8	0	**0,004331**
IFNA17	c.170dupA	p.His57fs	7	0	**0,007447**
IFNA21	c.304C>A	p.Gln102Lys	3	0	0,085164
IFNA5	c.343C>T	p.Gln115Ter	1	0	0,336389
IFNA6	c.284A>G	p.Asp95Gly	1	0	0,065201
IFNA7	c.385G>T	p.Val129Leu	1	0	0,310639
ULK1	c.1355C>G	p.Thr452Ser	3	0	0,058013
ULK1	c.1976G>A	p.Arg659Gln	2	0	**0,000004**
ULK1	c.2114C>T	p.Ala705Val	1	0	0,131322
ULK2	c.724C>T	p.Pro242Ser	4	1	**0,004134**
ULK2	c.931C>T	p.Pro311Ser	1	0	**0,044093**

Table 19| Table shows the variants found in our cohort of patients among the 39 genes of the Regulation of autophagy pathway. In bold, the variants significantly associated with LOPD in our cohort. **NA**: not available; **NFE**, non-Finnish European.

22.5.8 Associated-LOPD variants in genes of the Apoptosis pathway

Among the 85 genes connected to the Apoptosis pathway, we found 45 variants in 28 genes and 18 reached statistical significance for an association with LOPD in our cohort (Table 20). Among the genes involved are present genes of the TNF receptor superfamily (*TNFRSF10A*, *TNFRSF10B*, *TNFRSF10D* and *TNFSF10*), the *ATM* gene that encodes for the ATM Serine/Threonine Kinase, and genes encoding for the calpains (*CAPN1* and *CAPN2*) that are calcium-activated neutral proteases ubiquitously expressed.

			Frequency in our cohort (n=30)		Fisher's exact test (p value)
Gene	**cDNA change**	**Amino acid change**	**N° Heterozygous cases**	**N° Homozygous cases**	**LOPD vs GnomAD NFE**
APAF1	c.2329G>A	p.Glu777Lys	3	0	**0,00516**
APAF1	c.2345A>C	p.Asn782Thr	1	0	0,48435
APAF1	c.2612C>T	p.Ser871Leu	2	0	**0,010263**
ATM	c.1810C>T	p.Pro604Ser	1	0	0,08921
ATM	c.2119T>C	p.Ser707Pro	1	0	0,509857
ATM	c.3161C>G	p.Pro1054Arg	1	0	1
ATM	c.4388T>G	p.Phe1463Cys	1	0	**0,031578**
ATM	c.146C>G	p.Ser49Cys	1	0	0,532475
ATM	c.5558A>T	p.Asp1853Val	2	0	0,063932
BID	c.166A>G	p.Ser56Gly	3	0	0,086787
BIRC7	c.560C>G	p.Ala187Gly	1	0	**0,010205**
CAPN1	c.1800T>A	p.Asp600Glu	1	0	**0,003285**
CAPN1	c.785C>A	p.Ala262Asp	3	0	**0,000006**
CAPN2	c.233_234insCACGGTAGGAAGCG	p.Pro78fs	1	0	NA
CAPN2	c.1009C>T	p.Arg337Cys	2	0	**0,000172**
CASP10	c.1208A>G	p.Tyr403Cys	2	0	NA
CASP8	c.2T>C	p.Met1Thr	1	0	0,520658
CASP9	c.371A>C	p.Glu124Ala	1	0	**0,008323**
CASP9	c.469A>G	p.Met157Val	1	0	0,078228
CASP9	c.532C>T	p.Arg178Cys	1	0	**0,00232**
CFLAR	c.523_523del	p.Val175fs	1	0	NA
CSF2RB	c.934G>A	p.Asp312Asn	2	0	0,342589
DFFA	c.493T>C	p.Cys165Arg	1	0	0,065241
ENDOG	c.198G>C	p.Lys66Asn	2	0	**0,029221**
ENDOG	c.734G>T	p.Arg245Leu	1	0	0,13592
ENDOG	c.737C>T	p.Thr246Ile	1	0	0,132711
FAS	c.580G>A	p.Glu194Lys	1	0	0,133152
IL1R1	c.1031C>T	p.Thr344Met	1	0	0,276714
IL3	c.159G>T	p.Leu53Phe	1	0	**0,005557**
IRAK2	c.1315C>G	p.Leu439Val	3	0	0,445405
IRAK2	c.1507C>A	p.Leu503Ile	3	0	**0,002368**
IRAK2	c.140C>A	p.Ser47Tyr	1	0	1
IRAK2	c.607G>A	p.Val203Met	1	0	0,105552
IRAK3	c.346T>G	p.Tyr116Asp	1	0	**0,002324**
IRAK4	c.1172G>A	p.Arg391His	1	0	0,573283
NTRK1	c.2339G>A	p.Arg780Gln	1	0	0,309012
PIK3CG	c.2569A>G	p.Thr857Ala	4	0	0,300439
RELA	c.871G>A	p.Asp291Asn	1	0	0,297209
RIPK1	c.1706C>T	p.Ala569Val	2	0	0,397771
TNFRSF10A	c.1232T>C	p.Met411Thr	1	0	**0,002319**
TNFRSF10B	c.598delT	p.Ser200fs	1	0	**0,015668**
TNFRSF10D	c.826C>T	p.Arg276Cys	1	0	**0,000929**
TNFRSF10D	c.827G>A	p.Arg276His	1	0	0,530469
TNFSF10	c.141C>A	p.Asp47Glu	1	0	**0,04588**
TRAF2	c.412G>A	p.Ala138Thr	2	0	**0,001265**

Table 20| Table shows the variants found in our cohort of patients among the 85 genes of the Apoptosis pathway. In bold, variants significantly associated with LOPD in our cohort. **NA**: not available; **NFE**, non-Finnish European.

Discussion

Amyotrophic Lateral Sclerosis

Amyotrophic lateral sclerosis is a rare neuromuscular disorder in which genetics is gaining even more attention every day, especially upon the advent of NGS. Despite ALS is genetically determined in 10% of cases, many familial mutations have been observed in sporadic cases, thus in cases that do not have a positive family history for the disease. One possible reason is that, since ALS is a disorder with onset in middle-late adulthood in most cases, it is not possible to investigate parents to understand if the genetic mutations have been inherited, or that small pedigree may hide family inheritance of pathogenic mutations, ending up in familial cases that appear to be sporadic. Moreover, it is not uncommon that ALS patients have family members that are affected by other NDs, like PD, AD, FTD or other neuromuscular diseases. Indeed, it is well known that NDs share many pathological molecular mechanisms, such as mitochondrial dysfunctions, accumulation of misfolded proteins and DNA instability. The molecular overlapping among NDs represents one possible reason at the basis of variable clinical manifestation in ALS patients, leading to challenging diagnosis.

Since the discovery of *SOD1* as the first gene implicated in ALS, many other genetic contributors have been identified, and their number is increasing every year. To date, *C9ORF72*, *TARDBP*, *FUS*, *SOD1* and *VCP* are the most commonly mutated genes in ALS, responsible for approximately 60%–70% of fALS and about 10% of apparently sALS cases, however up to 25 genes have been considered causative for ALS but the number is much higher if considering either genes that confer susceptibility for ALS and genes that have been clearly linked to other NDs but that have been described also in patients with an ALS phenotype (i.e. *SPG11*, *SPG4*). The association of genes across different NDs as well as the discovery of new genes were made possible by the advent of the NGS technology that allowed the possibility to screen several genes per many patients simultaneously, with limited costs. The introduction of NGS in both research and diagnostic routine, overtaken the employment of the classic Sanger sequencing that instead allows the genetic screening of one or few genes per patient in a single run. Indeed, nowadays most laboratories that perform genetic analysis apply the NGS technology, using panels of genes specific for the disease under study. However, it is becoming common that laboratories decide to perform WES analysis instead of panels of selected genes. This choice is because WES allows to simultaneously screen all the coding regions

of the genome, with an output of more than one hundred thousand variants per patient. This huge amount of data represents a precious resource not only to look for variants strictly related to the pathology of interest but also to search other pathogenic variants in other genes as well to look for genetic modifiers.

Given this evidence, the aim of this part of the Ph.D. project was to investigate the genetic architecture of a cohort of ALS patients, through the employment of the NGS technology. Specifically, we recorded i) pathogenic variants in the already known genes associated with ALS, ii) rare variants that confer susceptibility to the disease, iii) risk factor variants either within ALS genes and within other diseases such as FTD, HSP, PD and other neurodegenerative and neuromuscular disorders. This study is the result of our laboratory experience which started from the employment of customized gene panels to end up with the utilization of WES technology and virtual panels. For this reason, our cohort of 248 patients comprises three subgroups of 60, 107 and 81 patients respectively. While the last group has been investigated through WES, the first two have been analysed using customized panels of 97 and 172 genes respectively.

Upon NGS screening, variants were filtered according to frequency, effect on the protein and pathogenicity: Hence, we retained only variants with a MAF < 0,01 in the general population that potentially alter the protein sequence (i.e. coding, splicing and 5'-UTR variants) and that are considered pathogenic, likely pathogenic or with uncertain significance according to the ACMG criteria. Since NGS correctly detects only single nucleotide variations or small insertions and/or deletions, the HRE within the *C9ORF72* gene has been detected through a repeat-primed PCR and HRE above 30 repeats were classified as pathogenic.

Across the three gene panels, we identified 38 pathogenic variants and 31 likely pathogenic variants (16%) out of 423 variants retrieved. In literature, pathogenic variants in ALS patients without a family history for the disease have been detected in up to 10% of cases. However, this data is the result of NGS analysis confined only to clearly ALS-linked genes. In this Ph.D project, we instead decided to enlarge the number of genes analyzed also to genes related to diseases that have been reported to partially overlap with ALS, or that occur in concomitant with ALS, for instance FTD, HSP and PD. Thus, the higher percentage of pathogenic and likely pathogenic variants found in this study is the result of the enlargement of the genetic screening

adopted in this study. Further investigation will be certainly needed to unravel the role of pathogenic variants in "non-canonical" genes.

Looking at the pathogenic variants, 25 were in ALS causative genes, followed by 5 variants in the HSP and PD associated genes respectively and the remaining three variants were in genes linked to other NDs. Among the 25 pathogenic variants in ALS genes,16 (6,4%) were HREs within the *C9ORF72* gene, being the major genetic contributors in our cohort of patients. This data is in accordance with the frequency of C9ORF72 among sALS[3] as well as with the frequency in the Italian population which is estimated to be 7,4%.[321] The age of onset of these patients ranged from 46 to 70 years old, the site of onset was variable with patients presenting either bulbar or spinal onset, and also the phenotype was variable, indeed patients presented either classic, PLMN, bulbar and flail leg phenotypes. Three out of 16 expanded patients had cognitive impairment, while six other patients presented a concomitant FTD, confirming the key role of *C9ORF72* in the FTD-motor neuron disease *continuum*. Interestingly, one patient carrying the HRE harboured also another variant in *FUS*, classified as VUS.

One pathogenic variant (p.Asn87Ser, N86S) was found in *SOD1*, the second most frequent gene involved in ALS in the European population after *C9ORF72*. The N86S mutation was described in a patient with a spinal onset, a fail leg phenotype and no cognitive impairment.[322] The age of onset in our patient was 58 years old, which is not in line with the mean age at onset SOD1-mutated sALS patients which is 41.4 years, suggesting that other genetic variants may have acted as modifiers for the age of onset. Kuuluvainen and colleagues reported an example concerning the modulation of the penetrance of the A90V mutation in the *SOD1* gene, due to secondary mutations harboured by the patient.[323] Even if our patient carry another mutation than the A90V we cannot exclude a modulatory effect of other mutations in our patient on the age of onset. Data regarding the rate progression in our patient were not available, however the patient is still alive after 30 months after diagnosis, accordingly with the literature that reports a rapid disease progression and survival times shorter than 3 years for this mutation.[324]

FIG4 homozygous mutations cause a severe form of AR CMT4J, with early-onset and involvement of both sensory and motor neurons. However, heterozygous deleterious variants in the *FIG4* gene have first been implicated in both sALS and fALS in 2009.[66,68] The stop-gain Q888X variant identified in our study is very rare

considering that it was observed in heterozygosis only once in the European (non-Finnish) population and it is not described in the literature. Osmanovic and colleagues suggested that *FIG4* variants are not causative alone, but ALS patients may need to carry rare variants in multiple genes to develop the disease. Our patient was a carrier of another missense variant (c.617A>G, K206R), within the *PANK2* gene, which is associated with the HARP syndrome (MIM: 607236) and the Neurodegeneration with brain iron accumulation 1 (MIM: 234200). Both are neurodegenerative conditions in which clinical manifestations that overlap with ALS clinical signs are present, such as dysarthria, corticospinal signs and progressive dementia. It is obvious that functional studies will be needed to assess the synergic role of these two variants in the phenotype of our patients, however, this data supports the new concept of ALS as an oligogenic disease in which many genetic determinants with a small or mild effect concur in the pathogenesis of ALS.[68]

Two mutations in the *TARDBP* gene of two patients were also found, in two contig amino acids. Both mutations, the p.Gly294Val (G294V) and p.Gly295Ser (G295S) fall within the main mutational hotspot of the gene in exon six, both affecting the glycine-rich domain of TDP-43 protein. In 2009, Del Bo and colleagues identified p.Gly294Val and p.Gly295Ser mutations in sALS cases.[325] The G294V amino acid substitution was identified in an Italian woman who developed progressive lower limb weakness at the age of 73, while the G295S mutation was observed in a 67 years old Italian male who developed upper limb weakness with atrophy.[325] In the same year, Corrado and collaborators also reported the two mutations in two patients. The G294V in a patient with bulbar onset and signs of dementia that arose 3 years before the onset of the motor neuronsl phenotype, and the G295S mutation is an ALS patient with spinal onset and FTD. Accordingly in our study, the patient carrying the G294V was a woman with an age at onset of 66 years of age, a bulbar onset and an intermediate progression rate, while the patient with the G295S mutation was a man with an age at onset of 61 years of age, a spinal onset and a fast progression rate with also manifested FTD.

Among the variants classified as likely pathogenic in the ALS genes, we found a missense variant (G507D) within the *FUS* gene. This variant was described in 2 unrelated Italian patients with sALS and an age of onset of 50 and 70 years of age, however, no clinical data were available for the two patients.[326] The G507D variant was described also by Hewitt et al. in a sporadic patient with predominant LMN ALS

involving both the lower and upper limbs and with an onset at age 69 years, who died of respiratory failure 42 months after symptom onset.[327] Our patients carrying the G507D missense mutation was a man who displayed a PLMN ALS started at 40 years of age and with a very slow progression of about 15 years. Despite the common G507D mutation shared by the patients described above, the clinical onset, as well as the disease progression, were highly variable, suggesting that other factors, genetic or environmental, may have played a role in phenotypic modulation.

Another patient harboured the p.Gly148Ser (G147S) mutation in the *SOD1* gene. This amino acidic change, as well as other three mutations within the codon 147 of *SOD1* have been already reported in the literature, suggesting that this amino acid is crucial for the function of the protein and mutations in this site strongly correlate with ALS. Andersen and colleagues described a case with the G147R mutation and a patient harboring the G147D mutation.[328] Two families with the G147D mutation were identified in a series of French fALS cases. The age of onset ranged from 45 to 73 years, with spinal onset as the most prevalent site of onset; individuals had no cognitive impairment and disease duration ranged from 10 to 49 months.[329] Canosa and colleagues described the G147C in three siblings of an Italian family.[330] The age on onset ranged from 46 to 52 years and the disease progression ranged from 10 to 27 months in the two siblings, while the proband was still alive after 24 months, at the time of the manuscript publication. All three cases presented spinal onset. In this study, we found the G147S in a patient with an age of onset of 63 years, a classic phenotype with an intermediate disease progression rate. The G147S mutation was reported in a sporadic case with bulbar onset at the age of 56 years and death from respiratory failure in 8 months.[331]

Among the pathogenic and likely pathogenic in our ALS cohort, there were also 13 mutations in genes linked to HSP. ALS and HSP are motor neuron diseases sharing clinical, pathological, and genetic similarities. Very recently, Osmanovic and colleagues reported the presence of heterozygous mutations in 4% of a cohort of 214 European ALS patients, proposing *SPG7* as a genetic risk factor for ALS.[332] Accordingly, we found five patients with heterozygous mutations in the SPG7 gene, while one patient was homozygous. This last patient had an age at onset of 58 years of age with spinal onset and pyramidal signs. Accordingly to the genotype, the clinical manifestation of the patients turned into an HSP. Concerning the patients with heterozygous mutations in the *SPG7* gene, the age of onset ranged from 68 to 78

years of age with a predominant slow progression and PLMN or bulbar phenotype. Only one patient displayed a fast progression, while only one patient had an early age of onset (40 years of age) however this patient was the carrier of the G507D in *FUS*, thus the age of onset may be related to this latter mutation. Our study confirms the findings of Osmanovic, reputing *SPG7* as a genetic risk factor for ALS. Beyond mutations in SPG7, other variants in genes linked to HSP have been found, involving for instance *ALDH18A1, CYP2U1,* and *SPG11* genes. The nucleotide substitution c.1393G>A in the *ALDH18A1* gene, causing the amino acidic change p.Glu465Lys (E465K), was found in heterozygosis. Generally, dominant and recessive mutations in *ALDH18A1* have been linked to HSP, causing dominant (SPG9A) and recessive (SPG9B) forms.[333] The comorbidity of HSP and ALS has been already described by Koh and colleagues, which reported a family of SPG9B patients, having a history of ALS and characterized by upper and lower motor neuron degeneration in the bulbar region and spinal regions.[334] Despite this, in the literature, publications describing the implications related to aforesaid mutation are not present. *CYP2U1* gene, encoding a member of the cytochrome P450 family 2, is associated with SPG56, a rare AR early-onset complicated form of HSP.[335] Up to now, there are no studies that directly correlate mutations within *CYP2U1* gene and the risk to develop ALS, hence further investigations will be needed to study more in-depth the stop-gained mutation observed in our patient. Generally, mutations in homozygosis within the *CYP2U1* gene are the genetic cause of early-onset forms of HSP; nonetheless, our adult patient exhibited a stop-gained variant in heterozygosis. Hence, over time this variant could be involved in the late-onset manifestation of the motoneuronal pathology that characterizes our patient. As in the case of the two patients with mutations in the *ALDH18A1* and the *CYP2U1* genes, the age of onset was delayed compared to the mean age of onset of ALS patients and were characterized by a slow progression.

SPG11 is another gene that connects HSP with ALS, indeed it has been linked to AR-HSP

and to juvenile ALS in homozygous state. We found the heterozygous c.5986dupT variant in the *SPG11* gene of one patient. The duplication causes a frameshift starting with codon Cysteine 1996 which turns into a Leucine, and creates a premature stop codon at position 4 of the new reading frame (p.Cys1996LeufsX4). This variant is predicted to cause loss of normal protein function either through protein truncation or nonsense-mediated mRNA decay. Furthermore, the c.5986dupT

variant is not observed at a significant frequency in large population cohorts.[336] The c.5986dupT variant in the *SPG11* gene has been reported previously in three siblings with gait difficulties, weakness, dysarthria, swallowing difficulties, and cognitive decline who also had a second *SPG11* pathogenic variant.[337] Lastly, considering that one of the purposes of this research was to study the genetic overlapping of the different neurodegenerative disorders among patients, this result could represent an important piece for the investigation. Further examinations will be needed to reveal the role and the implications of the cited mutations and how these variants, in association with the VUS identified, contribute to the ALS phenotype.

Our analysis allowed the identification also of ten mutations (five pathogenic and five likely pathogenic) in genes known to be causative of PD, mostly within *GBA*. Three out of four patients carried the p.Leu483Pro (L444P) missense mutation, while the other patient harboured the p.Thr449Met (T410M). More precisely, the L444P mutation was found in two patients as a principal pathogenic variant, while in the third patient was found as an additional mutation to the above discussed E465K affecting *ALDH18A1. GBA* constitutes the major genetic risk factor for PD; mutations within this gene account for 2–30% of PD patients, including L444P which is one of the most common variants worldwide.[338,339] This mutation is associated with an early onset, as well as a more rapid disease progression and increased involvement of cognitive functions among PD patients; moreover, it was also identified in patients affected by Dementia with Lewy Bodies.[338] Beyond mutations in the *GBA* gene, we found two patients with the same amino acid change in the *PRKN* gene (R275W). *PRKN* is one of the most commonly mutated recessive gene in PD, whose loss of function is associated with juvenile Parkinsonism and early-onset PD.[340] Up to now, it has not been reported a direct correlation between these aforesaid mutations and the risk to develop ALS, therefore functional analysis will be taken into consideration; even so, PD may associate with MND in various neurodegenerative conditions, such as multiple system atrophy, spinocerebellar ataxia (SCA), HSP, FTD and ALS.[341] The comorbidity of ALS and parkinsonism had been described in 1973;[342] nowadays, symptoms and signs of parkinsonism have been described in ALS patients, with a frequency ranging from 5% to 17%.[341] However, in our cohort of ALS cases, most of the patients carrying pathogenic or likely pathogenic variants in genes associated with PD displayed a classic phenotype with no extrapyramidal signs. Only one patient presented a parkinsonian syndrome. This patient harbored a heterozygous missense

mutation in *PINK1* (L268V). This variant was found in heterozygous state in a patient with early-onset PD,[343] and the age of onset of our patient was 32, suggesting that probably the ALS phenotype in this case may be a concomitant clinical manifestation beside PD. Indeed, it is not uncommon that stand-alone heterozygous mutations in *PINK1* have been found in PD patients.[344]

Eventually, through this genetic analysis, we identified other twelve variants in genes associated with other NDs, including *ADAR* (dyschromatosis symmetrical hereditaria), *NPC1* (Niemann-Pick type C disease), *GLB1* (GM1-gangliosidosis) and *HEXA* (GM2 gangliosidosis). The two mutations observed in *ADAR* and *NPC1* respectively, have never been described in the literature; furthermore, these two genes have not yet been associated with ALS. Actually, *NPC1* is linked to Niemann-Pick type C disease, an AR lipid storage disorder characterized by progressive neurodegeneration;[345] while *ADAR* is responsible for dyschromatosis symmetrical hereditaria, a rare pigmentary genodermatosis.[346] Mutations in *GLB1* gene result in GM1-gangliosidosis, an AR neurological disorder characterized by the accumulation of ganglioside substrates in lysosomes. Kumar *et al.* identified compound heterozygous variants in the *GLB1* gene in HSP patients, suggesting that GM1 gangliosidosis may be a phenocopy for a severe, early-onset, complicated form of HSP.[347] *HEXA* gene encodes for Hexosaminidase A (Hex-A), a lysosomal enzyme that participates in the degradation of the ganglioside GM2. Enzyme deficiency causes the accumulation of this ganglioside, leading to the degeneration of nerve cells and a wide spectrum of neurological diseases.[348] Total Hex-A deficiency is responsible for a fatal infantile disorder, the Tay-Sachs disease; whereas partial deficiency is associated with adult-onset neurological phenotypes, characterized by UMN and LMN signs, parkinsonism, cerebellar disturbances, psychosis or dementia in different combinations.[348] Previous articles highlighted that LMN symptom dominated in late-onset GM2 gangliosidosis, but cases mimicking PLS, ALS, FTD, Friedreich ataxia, corticobasal dementia or even supranuclear palsy were also reported.[349] Adult-onset GM2 gangliosidosis is commonly mentioned in the differential diagnosis of ALS, especially in atypical cases associated with other neurological disturbances;[348,350] however, a literature review revealed only a few ALS cases with Hex-A deficiencies. To determine if there is a possible relationship between ALS phenotype and *HEXA* mutations and to clarify whether an ALS-mimic syndrome is a manifestation of adult-onset GM2, Drory and collaborators screened a

cohort of sALS patients searching for the most prevalent *HEXA* mutations. Their study failed to reveal significant Hex-A deficiency among ALS patients and also led them to conclude that adult-type Hex-A deficiency is only rarely a potential cause for ALS-mimic syndromes.[348] For our study, further investigations will be needed to prove if the genetic contribution of *HEXA,* as well as of the other genes involved in other NDs, concur in the manifestation of an ALS phenotype.

Besides the 69 pathogenic and likely pathogenic mutations identified, we retrieved many VUS variants. As already explained previously, VUS represents a major challenge for genetic counseling and clinical management since their functional impact cannot be deduced from sequence information alone. However, they could have a synergistic role in eliciting the disease. As a matter of fact, the occurrence of mutations in two or more genes, which are together required to cause an ALS phenotype, could justify the considerable proportion of apparently familial cases with reduced penetrance as well as a substantial proportion of sporadic cases.[351]. Furthermore, it will be intriguing to reveal and evaluate the genetic contribution of VUS variants in the modulation of ALS phenotypes.

In the last years, the dichotomized genetic "mutated or not-mutated" concept regarding ALS has been replaced by the new vision of ALS as a polygenic disease. Hence, ALS may be caused by multiple rare variants with additive or synergistic effects on disease presentation.[352] These variants can interact with environmental factors to let ALS manifest or may indicate an oligogenic disease model where an ALS mutation of large effect, for instance variant within the *SOD1* or the *C9ORF72* genes, is inherited with another ALS gene variant and their summative effect may contribute to the ALS phenotype.[133,353] Accordingly, through this genetic analysis we also found patients with a mutation of large effect, such as *C9ORF72* accompanied by other mutation in canonical ALS genes, such as *FUS*, or again, we reported a patient with a well-known *FUS* mutation which was accompanied by a mutation in the *SPG7* gene, suggesting oligogenic bases for these patients.

Through this study, our aim was to highlight the amplitude of the phenotypic and the genetic diversity among Italian patients affected by ALS. To further investigate the genetic architecture of our cohort of cases, we collected and listed 726-implicate ALS variants within 25 genes associated with ALS and we sought genetic associations between the variants and our cohort of patients compared to the general population.

We established a high degree of genetic heterogeneity among our patients, with 26 different variants from 14 ALS genes identified among 54,4% of cases. Among the 26 variants, two were significantly associated with our cohort of ALS patients, an in-frame deletion (p.Gly173_Gly174del) within *FUS* and a missense variant (R208W) within *SIGMAR1*. As our cohort is mainly composed of sporadic ALS cases, thus with no known family history for the disease, this high prevalence of reported ALS-implicated variants may reflect unrecognized false positive variants in the literature. Indeed, the evidence supporting the pathogenicity of each ALS-implicated variant varies significantly. While many have strong support as pathogenic ALS mutations through genetic linkage studies and/or extensive segregation within families, others have less compelling evidence to support their role in ALS pathogenesis.[301] Additionally, 7 patients (3%) of our cohort presented more than one of the listed ALS-implicated variants. Also in this case, an hypothesis of oligogenic bases in which the clinical phenotype of the patients is likely influenced by the sum of the multiple ALS-implicated variants carried may be assumed.

In conclusion, through this Ph.D. project we depicted the heterogeneity of the genetic architecture of ALS, made up not only of mutation in canonical ALS genes, but also of the interplay between variants of different NDs like HSP, FTD, PD but also of other motor neuron and neurodegenerative disorders. Moreover, through the study of our cohort of patients using a comprehensive survey of implicated ALS-implicated variants, we support the idea of ALS as a polygenic disorder, in which clinical manifestations are influenced by the presence of multiple ALS-implicated variants. Future studies are likely to continue to uncover novel genetic susceptibility and modifier variants that influence the development, presentation and course of ALS.

Late-onset Pompe disease

Late-onset Pompe disease is a rare autosomal recessive disorder caused by mutations within the *GAA* gene, a lysosomal enzyme. Mutations in this gene lead to GAA enzyme deficiency and ultimately to glycogen accumulation in multiple tissues, especially in skeletal muscles. The phenotypic manifestation of LOPD patients is characterized by high variability concerning age of onset, severity and muscular and/or respiratory involvement. Indeed, patients affected by LOPD may have an age of onset that can range between one year of age and late-adulthood and may have different muscular involvement with different degrees of severity. It is not uncommon to see families in which siblings with the same genotype manifest different clinical manifestation, i.e. if an individual requires respiratory assistance and does not need support for the ambulation the other sibling may require wheelchair. Since LOPD is characterized by the loss of GAA enzymatic activity, to date the gold standard clinical examination for LOPD diagnosis is the evaluation of the enzymatic activity on dried blood spot. This analysis can be subsequently followed by genetic confirmation. However, from the moment that LOPD is a monogenic disorder, mostly based on biallelic point mutations, almost all genetic analyses are based only on the classic Sanger sequencing. However, through this method possible rearrangements or gross insertions/deletions may be not detected. The first part of this Ph.D. project aimed to identify possible deletions or duplications in a subset of patients diagnosed with LOPD in which only one pathogenic mutation has been observed. We found a heterozygous deletion of the first exon of the *GAA* gene in one patient, and the deletion was then inherited by a son. To our knowledge, except for the few cases of IOPD patients characterized by the deletion of a large portion or of the entire *GAA* gene,[354,355] Δ18 is the only large deletion associated to LOPD. The first exon of the *GAA* gene is a non-coding exon and it takes part of the 5' untranslated region (5'-UTR) that, together with the promoter region, is involved in the transcriptional regulation of the gene. Actually, the *GAA* gene shares the common characteristics of the housekeeping genes that are characterized by a first non-coding exon kilobases far from the second coding one.[356] Accordingly, within this exon and in its very proximity, many transcription sites are present, such as sites for the Sp-1 transcription factor, but also of many other transcriptional regulators. Therefore, the deletion of the first exon, which is a portion of the 5'-UTR, may prevent the binding of transcription factors and alter the transcriptional regulation of the *GAA* gene, and

thus of the mRNA levels. The amount of the *GAA* transcript was lower in the proband who harboured also the c.-32-13T>G, which is a variant known to alter mRNA levels, with respect a healthy control. Thus, the deletion in the first exon of the *GAA* gene, in combination with the c.-32-13T>G mutation, probably causes a pathological deficiency of the enzyme activity. Moreover, also the son of the proband who carried only of the deletion of the first exon of the gene displayed similar amounts of mRNA levels, reinforcing the hypothesis that this deletion may alter the transcriptional levels of the *GAA* gene. Further studies will be needed to characterize the exon one deletion in order to identify its breakpoint and thus to understand its real impact in the disruption of the 5'-UTR. Moreover, studies to investigate which transcription factors bind the deleted portion of the *GAA* gene and their impact of mRNA levels will be performed.

In the 9 cases in whom the clinical diagnosis of LOPD was followed by the identification of only one pathogenic mutation after both sequencing and MLPA analyses, several hypotheses can be discussed. Certainly, intronic sequence variations introducing new splicing sites is a possibility, but its likelihood is extremely low. Therefore, the possibility of a casual identification of a *GAA* mutation in a patient with a myopathy other than Pompe disease has a chance to be observed about 1 in 50.[357] Accordingly, in a recent paper by Savarese et al., biallelic *GAA* mutations were identified in 10/275 cases of patients with myopathy without a defined diagnosis, thus demonstrating that difficulties in clinical differential diagnosis are present in a consistent number of patients with myopathy.[358] Eventually, heterozygous *GAA* mutations, at least in theory, may add to the clinical expression of other genetic variations in genes relate to neuromuscular disorders. With the classical Sanger-based genetic testing, it is common to undetected heterozygous large genomic deletions due to the amplification of the undeleted allele. For that reason and on the bases of our findings, we strongly recommend to taking into consideration MLPA analysis as genetic test for LOPD diagnosis to detect potential deleterious loss of heterozygosity.

To date, a strong correlation between the severity of the clinical phenotype and the severity of the mutations harboured by the patients has been assumed. Indeed, Pompe disease manifestation relies on the residual activity of the GAA enzyme relies which ultimately depends on the type of mutation occurred. As a consequence, LOPD appears when the residual activity of the enzymes is lower than 30%, while

the more severe infantile form of PD occurs with no or less that 1% of enzymatic activity.[260,359] However, this statement does not explain the differences in the clinical phenotype of patients sharing the same genotype. To date, in the literature, no data concerning a correlation or an influence of environmental factors in the onset of LOPD has been reported, thus we hypothesized that genetic modifiers may have a role in the clinical modulation of the disease. Considering this, the second part of this project was aimed to perform WES analysis in a selected cohort of 30 LOPD patients to identify variants associated to our cohort of cases with respect the general population. Since WES allows to obtain hundreds of thousands of variants, we filtered variants in order to obtain only rare nucleotide substitutions with a MAF of 0.05 and with an effect on the protein sequence. We further created six virtual panels of genes related to the principal pathways involved in the pathogenesis of LOPD, for a total of 302 genes and we focused on the variants found in these genes within our cohort of patients. We found a total of 176 variants in 103 (34%) out of 302 genes analysed across the six virtual gene panels. The majority resided in genes linked to the Lysosomal pathway, followed by variants in genes associated with the Apoptosis and the Regulation of Autophagy pathways. In these three pathways, we retrieved respectively 36, 18 and 15 variants that resulted significantly associated to our cohort of LOPD patients with respect the general population. Beyond the Lysosomal pathway that is clearly involved in the aetiology of LOPD since the GAA enzyme is a lysosomal enzyme, Autophagy and its regulation and Apoptosis are two understudied pathways involved in the pathology of LOPD. Autophagy is a lysosomal-dependent process aimed at the degradation of long-lived proteins and damaged organelles. It is considered an important feature of the skeletal muscle pathology in patients with LOPD where the accumulation of glycogen in skeletal muscle is not limited to lysosomes but is also found in autophagic vacuoles with cytoplasmic degradation products.[360] This suggested the autophagic pathway as a potential therapeutic target. Additionally, in murine model skeletal muscle type II fibres showed an enhanced autophagy in comparison to type I fibres and resulted more resistant to enzymatic therapy.[361,362] Hence, the need for the suppression of autophagy in Pompe skeletal muscle has two reasons: i) to prevent the accumulation of a disruptive autophagic build-up in myofibers and ii) to reduce the glycogen load.[363] This data suggested that patients poor responders to ERT (DD45, FM63, GAA81) may be carrier of significantly implicated variants in LOPD in one the genes of the Regulation of

autophagy pathway. Unfortunately, among the variants significantly associated with our cohort there were not variants enriched in the pool of patients with poor response to the ERT therapy. Despite our results were inconclusive in finding genetic modifiers for ERT therapy, suppression of autophagy should be considered as a novel therapeutic approach that may be useful in diseases with disturbed autophagy like LOPD. Apoptosis, which is the programmed cell death, is also been considered as a mechanism involved in Pompe disease. Indeed the skeletal muscle wasting and the consequent loss of force that occur in LOPD patients, has been considered as a result of both disuse atrophy and apoptosis.[319] Hence, we verified if patients with the worse six minute walking test (RA01, SM01, FM63 and GAA81) were carrier of significantly associated variants with LOPD, compared to the three heathy subjects (SM02, SM04 and BR04). Also in this case, we did not found variants strictly related to one subgroup in relation to the clinical feature taken into consideration. The lack of variants associated with a specific clinical feature in our cohort of cases may be due to the small number of cases analysed, but also to the pathway taken into considerations, suggesting that more cross-analysis between pathways and clinical features should be done. Moreover, other pathway should be taken into account to identify possible genetic variants implicated in phenotype modulation in the context of LOPD.

Conclusion remarks

This experimental thesis has reaffirmed the heterogeneity of the genetic of ALS, which takes into consideration not only ALS-causative genes and risk factors in genes notably associated with ALS, but consider also genes associated with other NDs such as FTD, HSP and PD, as well as gene related to other motor neuron diseases, highlighting the importance of genetic comorbidity studies. Additionally, this study emphasized the key role of next generation technologies such as WES to study the genetics architecture of ALS, either for the identification of pathogenic variants causative of the disease, but also to identify all the genetic factors that predispose to and may explain not only the “classic” clinical phenotypes, but especially more complex phenotypes. Furthermore, this Ph.D. project confirmed the new vision of ALS as a polygenic disease characterized by oligogenic inheritance, underlying again the importance to obtain a complete view of the genetic architecture of ALS patients through complex analysis such as the WES genetic screening.

Additionally, this Ph.D. thesis underlies also the importance in the utilization of WES technology also in genetically less complex diseases, such as LOPD, which is a monogenic disorder. In this case indeed, WES technology may be useful in finding genetic modifiers able to explain the clinical variability observed between patients, especially in those that share the same *GAA* genotype.

References

1. Taylor, J. P., Brown, R. H. & Cleveland, D. W. Decoding ALS: From genes to mechanism. *Nature* **539**, 197–206 (2016).

2. Rowland, L. P. How amyotrophic lateral sclerosis got its name: The clinical-pathologic genius of Jean-Martin Charcot. *Archives of Neurology* vol. 58 512–515 (2001).

3. Mejzini, R. *et al.* ALS Genetics, Mechanisms, and Therapeutics: Where Are We Now? *Front. Neurosci.* **13**, 1–27 (2019).

4. Grad, L. I., Rouleau, G. A., Ravits, J. & Cashman, N. R. Clinical Spectrum of Amyotrophic Lateral Sclerosis (ALS). *Cold Spring Harb. Perspect. Med.* **7**, 1–15 (2017).

5. Filippini, T. *et al.* Clinical and lifestyle factors and risk of amyotrophic lateral sclerosis: A population-based case-control study. *Int. J. Environ. Res. Public Health* **17**, 1–17 (2020).

6. Huisman, M. H. B. *et al.* Population based epidemiology of amyotrophic lateral sclerosis using capture-recapture methodology. *J. Neurol. Neurosurg. Psychiatry* **82**, 1165–1170 (2011).

7. Longinetti, E. & Fang, F. Epidemiology of amyotrophic lateral sclerosis: An update of recent literature. *Curr. Opin. Neurol.* **32**, 771–776 (2019).

8. Al-Chalabi, A. & Hardiman, O. The epidemiology of ALS: A conspiracy of genes, environment and time. *Nature Reviews Neurology* vol. 9 617–628 (2013).

9. Ellison, D., Love, S. & Chimelli, L. *Neuropathology: a reference text of CNS pathology*. (2012).

10. Roccatagliata, L., Bonzano, L., Mancardi, G., Canepa, C. & Caponnetto, C. Detection of motor cortex thinning and corticospinal tract involvement by quantitative MRI in amyotrophic lateral sclerosis. *Amyotroph. Lateral Scler.* **10**, 47–52 (2009).

11. Kassubek, J. *et al.* Global brain atrophy and corticospinal tract alterations in ALS, as investigated by voxel-based morphometry of 3-D MRI. *Amyotroph. Lateral Scler. Other Mot. Neuron Disord.* **6**, 213–220 (2005).

12. Dickson, D. W. & Weller, R. O. *Neurodegeneration: The Molecular Pathology of Dementia and Movement Disorders: Second Edition*. *Neurodegeneration: The Molecular Pathology of Dementia and Movement Disorders: Second Edition* (Wiley-Blackwell, 2011). doi:10.1002/9781444341256.

13. Stephens, B. *et al.* Widespread loss of neuronal populations in the spinal ventral horn in sporadic motor neuron disease. A morphometric study. *J. Neurol. Sci.* **244**, 41–58 (2006).

14. Piao, Y. S. *et al.* Neuropathology with clinical correlations of sporadic amyotrophic lateral sclerosis: 102 Autopsy cases examined between 1962 and 2000. *Brain Pathol.* **13**, 10–22 (2003).

15. Mizuno, Y., Fujita, Y., Takatama, M. & Okamoto, K. Peripherin partially localizes in Bunina bodies in amyotrophic lateral sclerosis. *J. Neurol. Sci.* **302**, 14–18 (2011).

16. Mizuno, Y. *et al.* Transferrin localizes in Bunina bodies in amyotrophic lateral sclerosis. *Acta Neuropathol.* **112**, 597–603 (2006).

17. Sasaki, S., Komori, T. & Iwata, M. Neuronal inclusions in sporadic motor neuron disease are negative for alpha-synuclein. *Neurosci. Lett.* **397**, 15–19 (2006).

18. Mizuno, Y. *et al.* Immunoreactivities of p62, an ubiqutin-binding protein, in the spinal anterior horn cells of patients with amyotrophic lateral sclerosis. *J. Neurol. Sci.* **249**, 13–18 (2006).

19. McGeer, P. L. & McGeer, E. G. Inflammatory processes in amyotrophic lateral sclerosis. *Muscle and Nerve* vol. 26 459–470 (2002).

20. Barbeito, L. H. *et al.* A role for astrocytes in motor neuron loss in amyotrophic lateral sclerosis. in *Brain Research Reviews* vol. 47 263–274 (Brain Res Brain Res Rev, 2004).

21. Arai, T. *et al.* TDP-43 is a component of ubiquitin-positive tau-negative inclusions in

frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Biochem. Biophys. Res. Commun.* **351**, 602–611 (2006).

22. Neumann, M. *et al.* Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science (80-.).* **314**, 130–133 (2006).

23. De Boer, E. M. J. *et al.* TDP-43 proteinopathies: A new wave of neurodegenerative diseases. *Journal of Neurology, Neurosurgery and Psychiatry* vol. 92 86–95 (2021).

24. Tan, R. H., Ke, Y. D., Ittner, L. M. & Halliday, G. M. ALS/FTLD: experimental models and reality. *Acta Neuropathologica* vol. 133 177–196 (2017).

25. Wilson, C. M., Grace, G. M., Munoz, D. G., He, B. P. & Strong, M. J. Cognitive impairment in sporadic ALS: A pathologic continuum underlying a multisystem disorder. *Neurology* **57**, 651–657 (2001).

26. Prudlo, J. *et al.* TDP-43 pathology and cognition in ALS. *Neurology* **87**, 1019–1023 (2016).

27. Gregory, J. M. *et al.* Executive, language and fluency dysfunction are markers of localised TDP-43 cerebral pathology in non-demented ALS. *J. Neurol. Neurosurg. Psychiatry* **91**, (2019).

28. van Es, M. A. *et al.* Amyotrophic lateral sclerosis. *Lancet* **390**, 2084–2098 (2017).

29. Al-Chalabi, A. & Lewis, C. M. Modelling the effects of penetrance and family size on rates of sporadic and familial disease. *Hum. Hered.* **71**, 281–8 (2011).

30. Hanby, M. F. *et al.* The risk to relatives of patients with sporadic amyotrophic lateral sclerosis. in *Brain* vol. 134 3451–3454 (Oxford University Press, 2011).

31. Johnston, C. A. *et al.* Amyotrophic lateral sclerosis in an urban setting: A population based study of inner city London [2]. *Journal of Neurology* vol. 253 1642–1643 (2006).

32. Turner, M. R. *et al.* Controversies and priorities in amyotrophic lateral sclerosis. *The Lancet Neurology* vol. 12 310–322 (2013).

33. van Doormaal, P. T. C. *et al.* The role of de novo mutations in the development of amyotrophic lateral sclerosis. *Hum. Mutat.* **38**, 1534–1541 (2017).

34. Rosen, D. R. *et al.* Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature* **362**, 59–62 (1993).

35. Al-Chalabi, A. & Hardiman, O. The epidemiology of ALS: A conspiracy of genes, environment and time. *Nature Reviews Neurology* vol. 9 617–628 (2013).

36. Mathis, S., Goizet, C., Soulages, A. & Vallat, J. Journal of the Neurological Sciences Genetics of amyotrophic lateral sclerosis : A review. **399**, 217–226 (2019).

37. Hardiman, O. *et al.* Amyotrophic lateral sclerosis. *Nat. Rev. Dis. Prim.* **3**, (2017).

38. Brown, R. H. & Al-Chalabi, A. Amyotrophic Lateral Sclerosis. *N. Engl. J. Med.* **377**, 162–172 (2017).

39. Kaur, S. J., McKeown, S. R. & Rashid, S. Mutant SOD1 mediated pathogenesis of Amyotrophic Lateral Sclerosis. *Gene* **577**, 109–118 (2016).

40. Pansarasa, O. *et al.* Sod1 in amyotrophic lateral sclerosis: “ambivalent” behavior connected to the disease. *Int. J. Mol. Sci.* **19**, 1–13 (2018).

41. Pansarasa, O. *et al.* Sod1 in amyotrophic lateral sclerosis: “ambivalent” behavior connected to the disease. *International Journal of Molecular Sciences* vol. 19 (2018).

42. Sirangelo, I. & Iannuzzi, C. The role of metal binding in the amyotrophic lateral sclerosis-related aggregation of copper-zinc superoxide dismutase. *Molecules* vol. 22 (2017).

43. Haase, G. & Rabouille, C. Golgi fragmentation in ALS motor neurons. New mechanisms targeting microtubules, tethers, and transport vesicles. *Front. Neurosci.* **9**, 448 (2015).

44. Benkler, C. *et al.* Aggregated SOD1 causes selective death of cultured human motor neurons. *Sci. Rep.* **8**, (2018).

45. Tang, L., Ma, Y., Liu, X. L., Chen, L. & Fan, D. S. Better survival in female SOD1-mutant patients with ALS: A study of SOD1-related natural history. *Transl. Neurodegener.* **8**, (2019).

46. Luisa Conforti, F. *et al.* Further evidence that D90A-SOD1 mutation is recessively inherited in ALS patients in Italy. *Amyotroph. Lateral Scler.* **10**, 58–60 (2009).

47. Rosen, D. R. *et al.* A frequent ala 4 to val superoxide dismutase-1 mutation is associated with a rapidly progressive familial amyotrophic lateral sclerosis. *Hum. Mol. Genet.* **3**, 981–987 (1994).

48. Deshpande, D. *et al.* Synaptic FUS localization during motoneuron development and its accumulation in human ALS synapses. *Front. Cell. Neurosci.* **13**, 1–17 (2019).

49. Baradaran-Heravi, Y., Van Broeckhoven, C. & van der Zee, J. Stress granule mediated protein aggregation and underlying gene defects in the FTD-ALS spectrum. *Neurobiology of Disease* vol. 134 (2020).

50. Gromicho, M., Oliveira Santos, M., Pinto, A., Pronto-Laborinho, A. & De Carvalho, M. Young-onset rapidly progressive ALS associated with heterozygous FUS mutation. *Amyotroph. Lateral Scler. Front. Degener.* **18**, 451–453 (2017).

51. Imperatore, J. A., McAninch, D. S., Valdez-Sinon, A. N., Bassell, G. J. & Mihailescu, M. R. FUS Recognizes G Quadruplex Structures Within Neuronal mRNAs. *Front. Mol. Biosci.* **7**, (2020).

52. Notaro, A. *et al.* ALS-linked FUS mutations confer loss and gain of function in the nucleus by promoting excessive formation of dysfunctional paraspeckles. *Acta Neuropathol. Commun.* **7**, 7 (2019).

53. Bäumer, D. *et al.* Juvenile ALS with basophilic inclusions is a FUS proteinopathy with FUS mutations. *Neurology* **75**, 611–618 (2010).

54. Kwiatkowski, T. J. *et al.* Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science (80-.).* **323**, 1205–1208 (2009).

55. Vance, C. *et al.* Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science (80-.).* **323**, 1208–1211 (2009).

56. Rutherford, N. J. *et al.* Novel Mutations in TARDBP (TDP-43) in Patients with Familial Amyotrophic Lateral Sclerosis. **4**, 1–8 (2008).

57. Prasad, A., Bharathi, V., Sivalingam, V. & Girdhar, A. Molecular Mechanisms of TDP-43 Misfolding and Pathology in Amyotrophic Lateral Sclerosis. **12**, 1–36 (2019).

58. Kawakami, I., Arai, T. & Hasegawa, M. The basis of clinicopathological heterogeneity in TDP-43 proteinopathy. *Acta Neuropathol.* **138**, 751–770 (2019).

59. Smethurst, P. *et al.* Distinct responses of neurons and astrocytes to TDP-43 proteinopathy in amyotrophic lateral sclerosis. *Brain* **143**, 430–440 (2020).

60. Lattante, S., Rouleau, G. A. & Kabashi, E. TARDBP and FUS Mutations Associated with Amyotrophic Lateral Sclerosis: Summary and Update. *Hum. Mutat.* **34**, 812–826 (2013).

61. Chiò, A. *et al.* Large proportion of amyotrophic lateral sclerosis cases in sardinia due to a single founder mutation of the TARDBP gene. *Arch. Neurol.* **68**, 594–598 (2011).

62. Borghero, G. *et al.* A patient carrying a homozygous p.A382T TARDBP missense mutation shows a syndrome including ALS, extrapyramidal symptoms, and FTD. *Neurobiol. Aging* **32**, 2327.e1-2327.e5 (2011).

63. Kovacs, G. G. *et al.* TARDBP variation associated with frontotemporal dementia, supranuclear gaze palsy, and chorea . *Mov. Disord.* **24**, 1842–1847 (2009).

64. Borroni, B. *et al.* Mutation within TARDBP leads to frontotemporal dementia without motor neuron disease. *Hum. Mutat.* **30**, (2009).

65. Gagliardi, M. *et al.* Mutational analysis of TARDBP gene in patients affected by Parkinson's disease from Calabria. *J. Neurol. Sci.* **390**, 209–211 (2018).

66. Chow, C. Y. *et al.* Deleterious Variants of FIG4 , a Phosphoinositide Phosphatase , in Patients

with ALS. *Am. J. Hum. Genet.* **84**, 85–88 (2009).

67. Chow, C. Y. *et al.* Deleterious Variants of FIG4, a Phosphoinositide Phosphatase, in Patients with ALS. *Am. J. Hum. Genet.* **84**, 85–88 (2009).

68. Osmanovic, A. *et al.* FIG4 variants in central European patients with amyotrophic lateral sclerosis : a whole-exome and targeted sequencing study. 324–331 (2017) doi:10.1038/ejhg.2016.186.

69. Lenk, G. M. *et al.* Pathogenic mechanism of the FIG4 mutation responsible for charcot-marie-tooth disease CMT4J. *PLoS Genet.* **7**, (2011).

70. Osmanovic, A. *et al.* FIG4 variants in central European patients with amyotrophic lateral sclerosis: A whole-exome and targeted sequencing study. *Eur. J. Hum. Genet.* **25**, 324–331 (2017).

71. Chow, C. Y. *et al.* Mutation of FIG4 causes neurodegeneration in the pale tremor mouse and patients with CMT4J. *Nature* **448**, 68–72 (2007).

72. Markovinovic, A. *et al.* Optineurin in amyotrophic lateral sclerosis: Multifunctional adaptor protein at the crossroads of different neuroprotective mechanisms. *Prog. Neurobiol.* **154**, 1–20 (2017).

73. Maruyama, H. & Kawakami, H. Optineurin and amyotrophic lateral sclerosis. *Geriatr. Gerontol. Int.* **13**, 528–532 (2013).

74. Del Bo, R. *et al.* Novel optineurin mutations in patients with familial and sporadic amyotrophic lateral sclerosis. *J. Neurol. Neurosurg. Psychiatry* **82**, 1239–1243 (2011).

75. Feng, S. M. *et al.* Novel mutation in optineurin causing aggressive ALS+/−frontotemporal dementia. *Ann. Clin. Transl. Neurol.* **6**, 2377–2383 (2019).

76. Naruse, H. *et al.* Mutational analysis of familial and sporadic amyotrophic lateral sclerosis with OPTN mutations in Japanese population. *Amyotroph. Lateral Scler.* **13**, 562–566 (2012).

77. Tümer, Z. *et al.* A novel heterozygous nonsense mutation of the OPTN gene segregating in a Danish family with ALS. *Neurobiol. Aging* **33**, 208.e1-208.e5 (2012).

78. Weishaupt, J. H. *et al.* A novel optineurin truncating mutation and three glaucoma-associated missense variants in patients with familial amyotrophic lateral sclerosis in Germany. *Neurobiol. Aging* **34**, 1516.e9-1516.e15 (2013).

79. Van Blitterswijk, M. *et al.* Novel optineurin mutations in sporadic amyotrophic lateral sclerosis patients. *Neurobiol. Aging* **33**, 1016.e1-1016.e7 (2012).

80. Johnson, J. O. *et al.* Exome Sequencing Reveals VCP Mutations as a Cause of Familial ALS. *Neuron* **68**, 857–864 (2010).

81. Ludtmann, M. H. R. *et al.* Mutations in valosin-containing protein (VCP) decrease ADP/ATP translocation across the mitochondrial membrane and impair energy metabolism in human neurons. *J. Biol. Chem.* **292**, 8907–8917 (2017).

82. Clemen, C. S. *et al.* The heterozygous R155C VCP mutation: Toxic in humans! Harmless in mice? *Biochem. Biophys. Res. Commun.* **503**, 2770–2777 (2018).

83. Johnson, J. O. *et al.* Exome Sequencing Reveals VCP Mutations as a Cause of Familial ALS. *Neuron* **68**, 857–864 (2010).

84. Plewa, J. *et al.* A cross-sectional analysis of clinical evaluation in 35 individuals with mutations of the valosin-containing protein gene. *Neuromuscul. Disord.* **28**, 778–786 (2018).

85. Ling, S. C., Polymenidou, M. & Cleveland, D. W. Converging mechanisms in als and FTD: Disrupted RNA and protein homeostasis. *Neuron* vol. 79 416–438 (2013).

86. Gite, J., Milko, E., Brady, L. & Baker, S. K. Phenotypic convergence in Charcot-Marie-Tooth 2Y with novel VCP mutation. *Neuromuscul. Disord.* **30**, 232–235 (2020).

87. Gonzalez, M. A. *et al.* A novel mutation in VCP causes Charcot-Marie-Tooth Type 2 disease.

Brain **137**, 2897–2902 (2014).

88. Johnson, J. O. *et al.* Exome Sequencing Reveals VCP Mutations as a Cause of Familial ALS. *Neuron* **68**, 857–864 (2010).

89. DeJesus-Hernandez, M. *et al.* Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. *Neuron* **72**, 245–256 (2011).

90. Barker, H. V., Niblock, M., Lee, Y. B., Shaw, C. E. & Gallo, J. M. RNA misprocessing in C9orf72-linked neurodegeneration. *Frontiers in Cellular Neuroscience* vol. 11 (2017).

91. Frick, P. *et al.* Novel antibodies reveal presynaptic localization of C9orf72 protein and reduced protein levels in C9orf72 mutation carriers. *Acta Neuropathol. Commun.* **6**, 72 (2018).

92. Todd, T. W. & Petrucelli, L. Insights into the pathogenic mechanisms of Chromosome 9 open reading frame 72 (C9orf72) repeat expansions. *Journal of Neurochemistry* 145–162 (2016) doi:10.1111/jnc.13623.

93. Budini, M., Buratti, E., Morselli, E. & Criollo, A. Autophagy and its impact on neurodegenerative diseases: New roles for TDP-43 and C9orf72. *Frontiers in Molecular Neuroscience* vol. 10 (2017).

94. Renton, A. E. *et al.* A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* **72**, 257–268 (2011).

95. Moens, T. G., Partridge, L. & Isaacs, A. M. Genetic models of C9orf72: what is toxic? *Curr. Opin. Genet. Dev.* **44**, 92–101 (2017).

96. Freibaum, B. D. & Taylor, J. P. The role of dipeptide repeats in C9ORF72-related ALS-FTD. *Frontiers in Molecular Neuroscience* vol. 10 (2017).

97. Van Der Ende, E. L. *et al.* Unravelling the clinical spectrum and the role of repeat length in C9ORF72 repeat expansions. *Journal of Neurology, Neurosurgery and Psychiatry* (2021) doi:10.1136/jnnp-2020-325377.

98. Gijselinck, I. *et al.* The C9orf72 repeat size correlates with onset age of disease, DNA methylation and transcriptional downregulation of the promoter. *Mol. Psychiatry* **21**, 1112–1124 (2016).

99. Trojsi, F. *et al.* Comparative analysis of C9Orf72 and sporadic disease in a large multicenter ALS population: The effect of Male sex on survival of C9Orf72 positive patients. *Front. Neurosci.* **13**, 485 (2019).

100. van Es, M. A. *et al.* Amyotrophic lateral sclerosis. *Lancet* **390**, 2084–2098 (2017).

101. Le Gall, L. *et al.* Personalized Medicine Molecular and Cellular Mechanisms Affected in ALS. doi:10.3390/jpm10030101.

102. Crozat, A., Åman, P., Mandahl, N. & Ron, D. Fusion of CHOP to a novel RNA-binding protein in human myxoid liposarcoma. *Nature* **363**, 640–644 (1993).

103. Schwartz, J. C. *et al.* FUS binds the CTD of RNA polymerase II and regulates its phosphorylation at Ser2. *Genes Dev.* **26**, 2690–2695 (2012).

104. Buratti, E. & Baralle, F. E. Characterization and Functional Implications of the RNA Binding Properties of Nuclear Factor TDP-43, a Novel Splicing Regulator of CFTR Exon 9. *J. Biol. Chem.* **276**, 36337–36343 (2001).

105. Lagier-Tourenne, C. *et al.* Divergent roles of ALS-linked proteins FUS/TLS and TDP-43 intersect in processing long pre-mRNAs. *Nat. Neurosci.* **15**, 1488–1497 (2012).

106. Scotti, M. M. & Swanson, M. S. RNA mis-splicing in disease. *Nature Reviews Genetics* vol. 17 19–32 (2016).

107. Polymenidou, M. *et al.* Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. *Nat. Neurosci.* **14**, 459–468 (2011).

108. Seto, S. *et al.* Rab39a interacts with phosphatidylinositol 3-kinase and negatively regulates autophagy induced by lipopolysaccharide stimulation in macrophages. *PLoS One* **8**, (2013).

109. Chen, Y. Z. *et al.* DNA/RNA helicase gene mutations in a form of juvenile amyotrophic lateral sclerosis (ALS4). *Am. J. Hum. Genet.* **74**, 1128–1135 (2004).

110. Wu, D. *et al.* Angiogenin loss-of-function mutations in amyotrophic lateral sclerosis. *Ann. Neurol.* **62**, 609–617 (2007).

111. McBride, H. M., Neuspiel, M. & Wasiak, S. Mitochondria: More Than Just a Powerhouse. *Current Biology* vol. 16 (2006).

112. Herrero-Mendez, A. *et al.* The bioenergetic and antioxidant status of neurons is controlled by continuous degradation of a key glycolytic enzyme by APC/C-Cdh1. *Nat. Cell Biol.* **11**, 747–752 (2009).

113. Fiorito, V., Chiabrando, D. & Tolosano, E. Mitochondrial targeting in neurodegeneration: A heme perspective. *Pharmaceuticals* vol. 11 (2018).

114. Kawamata, H. & Manfredi, G. Mitochondrial dysfunction and intracellular calcium dysregulation in ALS. *Mech. Ageing Dev.* **131**, 517–526 (2010).

115. Crugnola, V. *et al.* Mitochondrial respiratory chain dysfunction in muscle from patients with amyotrophic lateral sclerosis. *Arch. Neurol.* **67**, 849–854 (2010).

116. Sotelo-Silveira, J. R. *et al.* Axonal mitochondrial clusters containing mutant SOD1 in transgenic models of ALS. *Antioxidants Redox Signal.* **11**, 1535–1545 (2009).

117. Bilsland, L. G. *et al.* Deficits in axonal transport precede ALS symptoms in vivo. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 20523–20528 (2010).

118. Boillée, S., Vande Velde, C. & Cleveland, D. W. W. ALS: A Disease of Motor Neurons and Their Nonneuronal Neighbors. *Neuron* vol. 52 39–59 (2006).

119. Hardiman, O. *et al.* Amyotrophic lateral sclerosis. *Nature Reviews Disease Primers* vol. 3 (2017).

120. Gupta, R. *et al.* The proline/arginine dipeptide from hexanucleotide repeat expanded C9ORF72 inhibits the proteasome. *eNeuro* **4**, (2017).

121. Dobrowolny, G. *et al.* Skeletal Muscle Is a Primary Target of SOD1G93A-Mediated Toxicity. *Cell Metab.* **8**, 425–436 (2008).

122. Magrané, J., Cortez, C., Gan, W. B. & Manfredi, G. Abnormal mitochondrial transport and morphology are common pathological denominators in SOD1 and TDP43 ALS mouse models. *Hum. Mol. Genet.* **23**, 1413–1424 (2014).

123. Deng, J. *et al.* FUS interacts with ATP synthase beta subunit and induces mitochondrial unfolded protein response in cellular and animal models. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E9678–E9686 (2018).

124. Deng, J. *et al.* FUS Interacts with HSP60 to Promote Mitochondrial Damage. *PLoS Genet.* **11**, (2015).

125. Chia, R., Chiò, A. & Traynor, B. J. Novel genes associated with amyotrophic lateral sclerosis: diagnostic and clinical implications. *The Lancet Neurology* vol. 17 94–102 (2018).

126. Juarez, J. C. *et al.* Superoxide dismutase 1 (SOD1) is essential for H2O 2-mediated oxidation and inactivation of phosphatases in growth factor signaling. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 7147–7152 (2008).

127. Smith, E. F., Shaw, P. J. & De Vos, K. J. The role of mitochondria in amyotrophic lateral sclerosis. *Neuroscience Letters* vol. 710 132933 (2019).

128. Chiò, A. *et al.* ATXN2 polyQ intermediate repeats are a modifier of ALS survival. *Neurology* **84**, 251–258 (2015).

129. Elden, A. C. *et al.* Ataxin-2 intermediate-length polyglutamine expansions are associated with

increased risk for ALS. *Nature* **466**, 1069–1075 (2010).

130. Bennett, C. L. *et al.* Senataxin mutations elicit motor neuron degeneration phenotypes and yield TDP-43 mislocalization in ALS4 mice and human patients. *Acta Neuropathol.* **136**, 425–443 (2018).

131. Higelin, J. *et al.* NEK1 loss-of-function mutation induces DNA damage accumulation in ALS patient-derived motoneurons. *Stem Cell Res.* **30**, 150–162 (2018).

132. Sukhanova, M. V., Singatulina, A. S., Pastré, D. & Lavrik, O. I. Fused in sarcoma (FUS) in dna repair: Tango with poly(ADP-ribose) polymerase 1 and compartmentalisation of damaged DNA. *International Journal of Molecular Sciences* vol. 21 1–18 (2020).

133. Van Rheenen, W. *et al.* Genome-wide association analyses identify new risk variants and the genetic architecture of amyotrophic lateral sclerosis. *Nat. Genet.* **48**, 1043–1048 (2016).

134. Sutton, M. A. & Schuman, E. M. Dendritic Protein Synthesis, Synaptic Plasticity, and Memory. *Cell* vol. 127 49–58 (2006).

135. Yuan, A., Rao, M. V., Veeranna & Nixon, R. A. Neurofilaments and neurofilament proteins in health and disease. *Cold Spring Harb. Perspect. Biol.* **9**, (2017).

136. Prokop, A. Cytoskeletal organization of axons in vertebrates and invertebrates. *Journal of Cell Biology* vol. 219 (2020).

137. Taylor, J. P., Brown, R. H. & Cleveland, D. W. Decoding ALS: From genes to mechanism. *Nature* vol. 539 197–206 (2016).

138. Fil, D. *et al.* Mutant Profilin1 transgenic mice recapitulate cardinal features of motor neuron disease. *Hum. Mol. Genet.* **26**, 686–701 (2017).

139. Wu, C. H. *et al.* Mutations in the profilin 1 gene cause familial amyotrophic lateral sclerosis. *Nature* **488**, 499–503 (2012).

140. Smith, B. N. *et al.* Exome-wide rare variant analysis identifies TUBA4A mutations associated with familial ALS. *Neuron* **84**, 324–331 (2014).

141. Puls, I. *et al.* Mutant dynactin in motor neuron disease. *Nat. Genet.* **33**, 455–456 (2003).

142. Nicolas, A. *et al.* Genome-wide Analyses Identify KIF5A as a Novel ALS Gene. *Neuron* **97**, 1268-1283.e6 (2018).

143. Netzahualcoyotzi, C. & Tapia, R. Degeneration of spinal motor neurons by chronic AMPA-induced excitotoxicity in vivo and protection by energy substrates. *Acta Neuropathol. Commun.* **3**, 27 (2015).

144. Wang, Y. & Qin, Z. H. Molecular and cellular mechanisms of excitotoxic neuronal death. *Apoptosis* **15**, 1382–1402 (2010).

145. Van Zundert, B. *et al.* Neonatal neuronal circuitry shows hyperexcitable disturbance in a mouse model of the adult-onset neurodegenerative disease amyotrophic lateral sclerosis. *J. Neurosci.* **28**, 10864–10874 (2008).

146. Shi, Y. *et al.* Haploinsufficiency leads to neurodegeneration in C9ORF72 ALS/FTD human induced motor neurons. *Nat. Med.* **24**, 313–325 (2018).

147. Selvaraj, B. T. *et al.* C9ORF72 repeat expansion causes vulnerability of motor neurons to Ca2+-permeable AMPA receptor-mediated excitotoxicity. *Nat. Commun.* **9**, (2018).

148. Shi, Y. *et al.* Identification and therapeutic rescue of autophagosome and glutamate receptor defects in C9ORF72 and sporadic ALS neurons. *JCI Insight* **4**, (2019).

149. Blasco, H., Mavel, S., Corcia, P. & Gordon, P. H. The Glutamate Hypothesis in ALS: Pathophysiology and Drug Development. *Curr. Med. Chem.* **21**, 3551–3575 (2014).

150. Petrov, D., Mansfield, C., Moussy, A. & Hermine, O. ALS clinical trials review: 20 years of failure. Are we any closer to registering a new treatment? *Frontiers in Aging Neuroscience* vol. 9 (2017).

151. Neumann, M. *et al.* Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science (80-.).* **314**, 130–133 (2006).

152. Butti, Z. & Patten, S. A. RNA dysregulation in amyotrophic lateral sclerosis. *Frontiers in Genetics* vol. 10 (2019).

153. Gill, C. *et al.* SOD1-positive aggregate accumulation in the CNS predicts slower disease progression and increased longevity in a mutant SOD1 mouse model of ALS. *Sci. Rep.* **9**, (2019).

154. Feneberg, E., Gray, E., Ansorge, O., Talbot, K. & Turner, M. R. Towards a TDP-43-Based Biomarker for ALS and FTLD. *Molecular Neurobiology* vol. 55 7789–7801 (2018).

155. De Boer, E. M. J. *et al.* TDP-43 proteinopathies: A new wave of neurodegenerative diseases. *Journal of Neurology, Neurosurgery and Psychiatry* vol. 92 86–95 (2021).

156. McAlary, L., Aquilina, J. A. & Yerbury, J. J. Susceptibility of mutant SOD1 to form a destabilized monomer predicts cellular aggregation and toxicity but not in vitro aggregation propensity. *Front. Neurosci.* **10**, (2016).

157. Deng, H. X. *et al.* Mutations in UBQLN2 cause dominant X-linked juvenile and adult-onset ALS and ALS/dementia. *Nature* **477**, 211–215 (2011).

158. Levy, J. R. *et al.* A motor neuron disease-associated mutation in p150Glued perturbs dynactin function and induces protein aggregation. *J. Cell Biol.* **172**, 733–745 (2006).

159. Ayaki, T. *et al.* Immunoreactivity of valosin-containing protein in sporadic amyotrophic lateral sclerosis and in a case of its novel mutant. *Acta Neuropathol. Commun.* **2**, (2014).

160. Johnson, J. O. *et al.* Mutations in the Matrin 3 gene cause familial amyotrophic lateral sclerosis. *Nat. Neurosci.* **17**, 664–666 (2014).

161. Balendra, R. & Isaacs, A. M. C9orf72-mediated ALS and FTD: multiple pathways to disease. *Nature Reviews Neurology* vol. 14 544–558 (2018).

162. Webster, C. P. *et al.* The C9orf72 protein interacts with Rab1a and the ULK 1 complex to regulate initiation of autophagy . *EMBO J.* **35**, 1656–1676 (2016).

163. Saeki, Y. JB special review - Recent topics in ubiquitin-proteasome system and autophagy: Ubiquitin recognition by the proteasome. *Journal of Biochemistry* vol. 161 113–124 (2017).

164. Katsuragi, Y., Ichimura, Y. & Komatsu, M. P62/SQSTM1 functions as a signaling hub and an autophagy adaptor. *FEBS Journal* vol. 282 4672–4678 (2015).

165. Paolicelli, R. C., Bergamini, G. & Rajendran, L. Cell-to-cell Communication by Extracellular Vesicles: Focus on Microglia. *Neuroscience* vol. 405 148–157 (2019).

166. Raposo, G. & Stoorvogel, W. Extracellular vesicles: Exosomes, microvesicles, and friends. *Journal of Cell Biology* vol. 200 373–383 (2013).

167. Silverman, J. M. *et al.* CNS-derived extracellular vesicles from superoxide dismutase 1 (SOD1)G93A ALS mice originate from astrocytes and neurons and carry misfolded SOD1. *J. Biol. Chem.* **294**, 3744–3759 (2019).

168. Sproviero, D. *et al.* Leukocyte derived microvesicles as disease progression biomarkers in slow progressing amyotrophic lateral sclerosis patients. *Front. Neurosci.* **13**, (2019).

169. Westergard, T. *et al.* Cell-to-Cell Transmission of Dipeptide Repeat Proteins Linked to C9orf72-ALS/FTD. *Cell Rep.* **17**, 645–652 (2016).

170. Kamelgarn, M. *et al.* Proteomic analysis of FUS interacting proteins provides insights into FUS function and its role in ALS. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1862**, 2004–2014 (2016).

171. Sackmann, C., Sackmann, V. & Hallbeck, M. TDP-43 Is Efficiently Transferred Between Neuron-Like Cells in a Manner Enhanced by Preservation of Its N-Terminus but Independent of Extracellular Vesicles. *Front. Neurosci.* **14**, (2020).

172. Scaramozza, A. *et al.* Skeletal muscle satellite cells in amyotrophic lateral sclerosis. *Ultrastruct.*

Pathol. **38**, 295–302 (2014).

173. Farg, M. A. *et al.* Erratum: C9ORF72, implicated in amytrophic lateral sclerosis and frontotemporal dementia, regulates endosomal trafficking [Human Molecular Genetics, 23, 13, (2014), (3579-3595)] DOI: 10.1093/hmg/ddu068. *Human Molecular Genetics* vol. 26 4093–4094 (2017).

174. O'Rourke, J. G. *et al.* C9orf72 is required for proper macrophage and microglial function in mice. *Science (80-.).* **351**, 1324–1329 (2016).

175. Aoki, Y. *et al.* C9orf72 and RAB7L1 regulate vesicle trafficking in amyotrophic lateral sclerosis and frontotemporal dementia. *Brain* **140**, 887–897 (2017).

176. Buratta, S. *et al.* Lysosomal exocytosis, exosome release and secretory autophagy: The autophagic- and endo-lysosomal systems go extracellular. *Int. J. Mol. Sci.* **21**, (2020).

177. Bug, M. & Meyer, H. Expanding into new markets - VCP/p97 in endocytosis and autophagy. *Journal of Structural Biology* vol. 179 78–82 (2012).

178. Martyn, C. & Li, J. Fig4 deficiency: A newly emerged lysosomal storage disorder? *Progress in Neurobiology* vols 101–102 35–45 (2013).

179. Chow, C. Y. *et al.* Deleterious Variants of FIG4, a Phosphoinositide Phosphatase, in Patients with ALS. *Am. J. Hum. Genet.* **84**, 85–88 (2009).

180. Pérez-Brangulí, F. *et al.* Dysfunction of spatacsin leads to axonal pathology in SPG11-linked hereditary spastic paraplegia. *Hum. Mol. Genet.* **23**, 4859–4874 (2014).

181. Ghavami, S. *et al.* Autophagy and apoptosis dysfunction in neurodegenerative disorders. *Prog. Neurobiol.* **112**, 24–49 (2014).

182. Gordon, P. H., Cheng, B., Katz, I. B., Mitsumoto, H. & Rowland, L. P. Clinical features that distinguish PLS, upper motor neuron-dominant ALS, and typical ALS. *Neurology* **72**, 1948–1952 (2009).

183. Tartaglia, M. C. *et al.* Differentiation between primary lateral sclerosis and amyotrophic lateral sclerosis: Examination of symptoms and signs at disease onset and during follow-up. *Arch. Neurol.* **64**, 232–236 (2007).

184. Wijesekera, L. C. *et al.* Natural history and clinical features of the flail arm and flail leg ALS variants. *Neurology* **72**, 1087–1094 (2009).

185. Lu, C. H. *et al.* Neurofilament light chain: A prognostic biomarker in amyotrophic lateral sclerosis. *Neurology* **84**, 2247–2257 (2015).

186. Ferguson, T. A. & Elman, L. B. Clinical presentation and diagnosis of Amyotrophic Lateral Sclerosis. *NeuroRehabilitation* **22**, 409–416 (2007).

187. Richards, D., Morren, J. A. & Pioro, E. P. Time to diagnosis and factors affecting diagnostic delay in amyotrophic lateral sclerosis. (2020) doi:10.1016/j.jns.2020.117054.

188. Benatar, M. *et al.* ALS biomarkers for therapy development: State of the field and future directions. *Muscle and Nerve* **53**, 169–182 (2016).

189. Turner, M. R. *et al.* Mechanisms , models and biomarkers in amyotrophic lateral sclerosis. **8421**, (2013).

190. Vejux, A., Namsi, A., Nury, T., Moreau, T. & Lizard, G. Biomarkers of amyotrophic lateral sclerosis: Current status and interest of oxysterols and phytosterols. *Front. Mol. Neurosci.* **11**, 1–13 (2018).

191. Gagliardi, D., Bresolin, N., Comi, G. Pietro & Corti, S. Extracellular vesicles and amyotrophic lateral sclerosis: from misfolded protein vehicles to promising clinical biomarkers. *Cell. Mol. Life Sci.* **1**, 3.

192. Sproviero, D. *et al.* Pathological proteins are transported by extracellular vesicles of sporadic amyotrophic lateral sclerosis patients. *Front. Neurosci.* **12**, (2018).

193. Ravnik-Glavač, M. & Glavač, D. Circulating RNAs as potential biomarkers in amyotrophic lateral sclerosis. *International Journal of Molecular Sciences* vol. 21 (2020).

194. Gagliardi, S. *et al.* Long non-coding and coding RNAs characterization in Peripheral Blood Mononuclear Cells and Spinal Cord from Amyotrophic Lateral Sclerosis patients. *Sci. Rep.* **8**, 2378 (2018).

195. De Aguilar, J. L. G. Lipid biomarkers for amyotrophic lateral sclerosis. *Front. Neurol.* **10**, (2019).

196. Bang, J., Spina, S. & Miller, B. L. Frontotemporal dementia. *Lancet* **386**, 1672–1682 (2015).

197. Coyle-Gilchrist, I. T. S. *et al.* Prevalence, characteristics, and survival of frontotemporal lobar degeneration syndromes. *Neurology* **86**, 1736–1743 (2016).

198. Pickering-Brown, S. M. *et al.* Inherited frontotemporal dementia in nine British families associated with intronic mutations in the tau gene. *Brain* **125**, 732–751 (2002).

199. Hogan, D. B. *et al.* The prevalence and incidence of frontotemporal dementia: A systematic review. *Canadian Journal of Neurological Sciences* vol. 43 S96–S109 (2016).

200. Gorno-Tempini, M. L. *et al.* Classification of primary progressive aphasia and its variants. *Neurology* **76**, 1006–1014 (2011).

201. Mesulam, M. Afasia Progressiva Primária: Uma Demência Da Rede De Linguagem. *Dementia e Neuropsychologia* vol. 7 2–9 (2013).

202. Bang, J., Spina, S. & Miller, B. L. Frontotemporal dementia. *The Lancet* vol. 386 1672–1682 (2015).

203. Arcand, M. End-of-life issues in advanced dementia part 2: Management of poor nutritional intake, dehydration, and pneumonia. *Canadian Family Physician* vol. 61 337 (2015).

204. Abramzon, Y. A., Fratta, P., Traynor, B. J. & Chia, R. The Overlapping Genetics of Amyotrophic Lateral Sclerosis and Frontotemporal Dementia. *Front. Neurosci.* **14**, 1–10 (2020).

205. Gozzano M. Sulla istopatologia e la patogenesi della sclerosi laterale amiotrofica. *Riv. Neurol.* **5**, 165–216 (1936).

206. De Caro D. Un caso di sclerosi laterale amiotrofica con demenza. Contributo allo studio delle lesioni corticali nella malattia di Charcot. *Rass. Stud. Psichiatr.* **30**, 705–722 (1942).

207. Burrell, J. R. *et al.* The frontotemporal dementia-motor neuron disease continuum. *The Lancet* vol. 388 919–931 (2016).

208. Ringholz, G. M. *et al.* Prevalence and patterns of cognitive impairment in sporadic ALS. *Neurology* **65**, 586–590 (2005).

209. S, Z., B, P., C, M., A, D. & V, S. Amyotrophic lateral sclerosis and frontotemporal dementia (ALS-FTD). *Arch. Ital. Biol.* **149**, (2011).

210. Atassi, N. *et al.* Ultra high-field (7tesla) magnetic resonance spectroscopy in Amyotrophic Lateral Sclerosis. *PLoS One* **12**, e0177680 (2017).

211. Buhour, M. S. *et al.* Voxel-based mapping of grey matter volume and glucose metabolism profiles in amyotrophic lateral sclerosis. *EJNMMI Res.* **7**, 21 (2017).

212. Laere, K. Van *et al.* Value of 18fluorodeoxyglucose-positron-emission tomography in amyotrophic lateral sclerosis a prospective study. *JAMA Neurol.* **71**, 553–561 (2014).

213. Morita, M. *et al.* A locus on chromosome 9p confers susceptibility to ALS and frontotemporal dementia. *Neurology* **66**, 839–844 (2006).

214. Vance, C. *et al.* Familial amyotrophic lateral sclerosis with frontotemporal dementia is linked to a locus on chromosome 9p13.2-21.3. *Brain* **129**, 868–876 (2006).

215. Shahheydari, H. *et al.* Protein quality control and the amyotrophic lateral sclerosis/frontotemporal dementia continuum. *Frontiers in Molecular Neuroscience* vol. 10 119

(2017).

216. de Souza, P. V. S., de Rezende Pinto, W. B. V., de Rezende Batistella, G. N., Bortholin, T. & Oliveira, A. S. B. Hereditary Spastic Paraplegia: Clinical and Genetic Hallmarks. *Cerebellum* **16**, 525–551 (2017).

217. Bellofatto, M., De Michele, G., Iovino, A., Filla, A. & Santorelli, F. M. Management of hereditary spastic paraplegia: A systematic review of the literature. *Front. Neurol.* **10**, 1–10 (2019).

218. Harding, A. E. CLASSIFICATION OF THE HEREDITARY ATAXIAS AND PARAPLEGIAS. *The Lancet* vol. 321 1151–1155 (1983).

219. Ruano, L., Melo, C., Silva, M. C. & Coutinho, P. The Global Epidemiology of Hereditary Ataxia and Spastic Paraplegia: A Systematic Review of Prevalence Studies. *Neuroepidemiology* **42**, 174–183 (2014).

220. Pensato, V. *et al.* Overlapping phenotypes in complex spastic paraplegias SPG11, SPG15, SPG35 and SPG48. *Brain* **137**, 1907–1920 (2014).

221. Orlacchio, A. *et al.* SPATACSIN mutations cause autosomal recessive juvenile amyotrophic lateral sclerosis. *Brain* **133**, 591–598 (2010).

222. Yang, Y. *et al.* The gene encoding alsin, a protein with three guanine-nucleotide exchange factor domains, is mutated in a form of recessive amyotrophic lateral sclerosis. *Nat. Genet.* **29**, 160–165 (2001).

223. Timmerman, V., Clowes, V. E. & Reid, E. Overlapping molecular pathological themes link Charcot-Marie-Tooth neuropathies and hereditary spastic paraplegias. *Exp. Neurol.* **246**, 14–25 (2013).

224. Meyer, T. *et al.* Early-onset ALS with long-term survival associated with spastin gene mutation. *Neurology* **65**, 141–143 (2005).

225. Brugman, F. *et al.* Spastin mutations in sporadic adult-onset upper motor neuron syndromes. *Ann. Neurol.* **58**, 865–869 (2005).

226. Boutry, M., Morais, S. & Stevanin, G. Update on the Genetics of Spastic Paraplegias. *Curr. Neurol. Neurosci. Rep.* **19**, (2019).

227. Boutry, M., Morais, S. & Stevanin, G. Update on the Genetics of Spastic Paraplegias. *Curr. Neurol. Neurosci. Rep.* **19**, (2019).

228. Thomas, B. & Beal, M. F. Molecular insights into Parkinson's disease. *F1000 Med. Rep.* **3**, 1–8 (2011).

229. FAHN, S. PDFlib PLOP : PDF Linearization , Optimization , Protection Page inserted by evaluation version Description of Parkinson ' s Disease as a. *New York Acad. Sci.* **991**, 1–14 (2003).

230. Shulman, J. M., De Jager, P. L. & Feany, M. B. Parkinson's disease: Genetics and pathogenesis. *Annu. Rev. Pathol. Mech. Dis.* **6**, 193–222 (2011).

231. Dawson, T. M. Parkinson's disease: Genetics and pathogenesis. *Park. Dis. Genet. Pathog.* 1–386 (2007) doi:10.1146/annurev-pathol-011110-130242.

232. Brait, K., Fahn, S. & Schwarz, G. A. Sporadic and familial parkinsonism and motor neuron disease. *Neurology* **23**, 990–1002 (1973).

233. Pupillo, E. *et al.* Extrapyramidal and cognitive signs in amyotrophic lateral sclerosis: A population based cross-sectional study. *Amyotroph. Lateral Scler. Front. Degener.* **16**, 324–330 (2015).

234. Pradat, P. F. *et al.* Extrapyramidal stiffness in patients with amyotrophic lateral sclerosis. *Mov. Disord.* **24**, 2143–2148 (2009).

235. Przedborski, S., Vila, M. & Jackson-Lewis, V. Neurodegeneration: What is it and where are we? *J. Clin. Invest.* **111**, 3–10 (2003).

236. Woulfe, J. Nuclear bodies in neurodegenerative disease. *Biochim. Biophys. Acta - Mol. Cell Res.* **1783**, 2195–2206 (2008).

237. Soto, C. & Pritzkow, S. Protein misfolding, aggregation, and conformational strains in neurodegenerative diseases. *Nat. Neurosci.* **21**, 1332–1340 (2018).

238. Matej, R., Tesar, A. & Rusina, R. Alzheimer's disease and other neurodegenerative dementias in comorbidity: A clinical and neuropathological overview. *Clin. Biochem.* **73**, 26–31 (2019).

239. Chiò, A. *et al.* The multistep hypothesis of ALS revisited. *Neurology* **91**, e635–e642 (2018).

240. Pompe JC. Over idiopathische hypertrophie van het hart. *Ned Tijdschr Geneeskd* **76**, 304–312 (1932).

241. Novikoff, A. B., Beaufay, H. & De Duve, C. Electron microscopy of lysosome-rich fractions from rat liver. *J. Cell Biol.* **2**, 179–184 (1956).

242. LEJEUNE, N., THINES-SEMPOUX, D. & HERS, H. G. Tissue fractionation studies. 16. Intracellular distribution and properties of alpha-glucosidases in rat liver. *Biochem. J.* **86**, 16–21 (1963).

243. Martiniuk, F. *et al.* Carrier frequency for glycogen storage disease type II in New York and estimates of affected individuals born with the disease [4]. *Am. J. Med. Genet.* **79**, 69–72 (1998).

244. Leslie, N. & Bailey, L. *Pompe Disease*. *GeneReviews*® (In: Adam MP, Ardinger HH, Pagon RA et al (eds) GeneReviews®.Seattle (WA). University of Washington, Seattle, pp 1993–2018, 2007).

245. Löscher, W. N. *et al.* Pompe disease in Austria: clinical, genetic and epidemiological aspects. *J. Neurol.* **265**, 159–164 (2018).

246. Morales, J. A. & Anilkumar, A. C. *Glycogen Storage Disease Type II (Pompe Disease)*. *StatPearls* (StatPearls Publishing, 2020).

247. Winkel, L. P. F. *et al.* The natural course of non-classic Pompe's disease; a review of 225 published cases. *Journal of Neurology* vol. 252 875–884 (2005).

248. Llerena Junior, J. C. *et al.* Diretriz para o diagnóstico, tratamento e acompanhamento clínico de pacientes com doença de Pompe juvenil e do adulto. *Arq. Neuropsiquiatr.* **74**, 166–176 (2016).

249. Hirschhorn, R. & Reuser, A. J. The Metabolic and Molecular Bases of Inherited Disease. in (ed. C.R. Scriver, A.L. Be audet, W.S. Sly, D. Valle, B. Vogelstein, B. C.) 3389–3420 (McGraw-Hill Publishing, 2001).

250. Martiniuk, F. *et al.* Isolation of a cDNA for human acid α-glucosidase and detection of genetic heterogeneity for mRNA in three α-glucosidase-deficient patients. *Proc. Natl. Acad. Sci. U. S. A.* **83**, 9641–9644 (1986).

251. Tzall, S. & Martiniuk, F. Identification of the promoter region and gene expression for human acid alpha glucosidase. *Biochem. Biophys. Res. Commun.* **176**, 1509–1515 (1991).

252. Wisselaar, H. A., Kroos, M. A., Hermans, M. P., Beeumens, J. Van & Reusers, A. J. J. Structural and Functional Changes of Lysosomal Acid a-Glucosidase during Intracellular Transport and Maturation *. **268**, 2223–2231 (1993).

253. Peruzzo, P., Pavan, E. & Dardis, A. Molecular genetics of Pompe disease: a comprehensive overview. *Ann. Transl. Med.* **7**, 278–278 (2019).

254. Laforêt, P. *et al.* Juvenile and adult-onset acid maltase deficiency in France: Genotype-phenotype correlation. *Neurology* **55**, 1122–1128 (2000).

255. Montalvo, A. L. E. *et al.* Mutation profile of the GAA gene in 40 Italian patients with late onset glycogen storage disease type II. *Hum. Mutat.* **27**, 999–1006 (2006).

256. Hule, M. L. *et al.* Aberrant splicing in adult onset glycogen storage disease type II (GSDII): Molecular identification of an IVS1 (- 13T→G) mutation in a majority of patients and a novel

IVS10 (+ 1GT → CT) mutation. *Hum. Mol. Genet.* **3**, 2231–2236 (1994).

257. Ausems, M. G. E. M. *et al.* Dutch patients with glycogen storage disease type II show common ancestry for the 525delT and del exon 18 mutations [1]. *Journal of Medical Genetics* vol. 38 527–529 (2001).

258. Hermans, M. M. P. *et al.* The effect of a single base pair deletion (δT525) and a C1634T missense mutation (pro545leu) on the expression of lysosomal α-glucosidase in patients with glycogen storage disease type II. *Hum. Mol. Genet.* **3**, 2213–2218 (1994).

259. Kroos, M. A. *et al.* Glycogen storage disease type II: Frequency of three common mutant alleles and their associated clinical phenotypes studied in 121 patients. *Journal of Medical Genetics* vol. 32 836–837 (1995).

260. Zampieri, S. *et al.* Splicing mutations in glycogen-storage disease type II: evaluation of the full spectrum of mutations and their relation to patients' phenotypes. *Eur. J. Hum. Genet.* **19**, 422–431 (2011).

261. Kroos, M. *et al.* Update of the Pompe disease mutation database with 107 sequence variants and a format for severity rating. *Hum. Mutat.* **29**, (2008).

262. Kroos, M. *et al.* Update of the pompe disease mutation database with 60 novel GAA sequence variants and additional studies on the functional effect of 34 previously reported variants. *Hum. Mutat.* **33**, 1161–1165 (2012).

263. Kroos, M. A. *et al.* Broad spectrum of Pompe disease in patients with the same c.-32-13T→G haplotype. *Neurology* **68**, 110–115 (2007).

264. Kroos, M., Hoogeveen-Westerveld, M., van der Ploeg, A. & Reuser, A. J. J. The genotype-phenotype correlation in Pompe disease. *Am. J. Med. Genet. Part C Semin. Med. Genet.* **160 C**, 59–68 (2012).

265. Wens, S. C. A. *et al.* Phenotypical variation within 22 families with Pompe disease. *Orphanet J. Rare Dis.* **8**, (2013).

266. De Filippi, P. *et al.* The angiotensin-converting enzyme insertion/deletion polymorphism modifies the clinical outcome in patients with Pompe disease. *Genet. Med.* **12**, 206–211 (2010).

267. De Filippi, P. *et al.* Genotype-phenotype correlation in Pompe disease, a step forward. *Orphanet J. Rare Dis.* **9**, (2014).

268. Baek, R. C. *et al.* The influence of a polymorphism in the gene encoding angiotensin converting enzyme (ACE) on treatment outcomes in late-onset Pompe patients receiving alglucosidase alfa. *Mol. Genet. Metab. Reports* **8**, 48–50 (2016).

269. Bergsma, A. J. *et al.* A genetic modifier of symptom onset in Pompe disease. *EBioMedicine* **43**, 553–561 (2019).

270. Thurberg, B. L. *et al.* Characterization of pre- and post-treatment pathology after enzyme replacement therapy for Pompe disease. *Lab. Investig.* **86**, 1208–1220 (2006).

271. Griffin, J. L. Infantile acid maltase deficiency - I. Muscle fiber destruction after lysosomal rupture. *Virchows Arch. B Cell Pathol. Incl. Mol. Pathol.* **45**, 23–36 (1984).

272. Kohler, L., Puertollano, R. & Raben, N. Pompe Disease: From Basic Science to Therapy. *Neurotherapeutics* vol. 15 928–942 (2018).

273. Li, W. W., Li, J. & Bao, J. K. Microautophagy: Lesser-known self-eating. *Cellular and Molecular Life Sciences* vol. 69 1125–1136 (2012).

274. Fukuda, T. *et al.* Autophagy and Mistargeting of Therapeutic Enzyme in Skeletal Muscle in Pompe Disease. *Mol. Ther.* **14**, 831–839 (2006).

275. Schoser, B. G. H. *et al.* Adult-onset glycogen storage disease type 2: Clinico-pathological phenotype revisited. *Neuropathol. Appl. Neurobiol.* **33**, 544–559 (2007).

276. van der Ploeg, A. T. *et al.* European consensus for starting and stopping enzyme replacement

therapy in adult patients with Pompe disease: a 10-year experience. *Eur. J. Neurol.* **24**, 768-e31 (2017).

277. van der Ploeg, A. T. & Reuser, A. J. Pompe's disease. *Lancet* **372**, 1342–1353 (2008).

278. Musumeci, O. *et al.* LOPED study: Looking for an early diagnosis in a late-onset Pompe disease high-risk population. *J. Neurol. Neurosurg. Psychiatry* **87**, 5–11 (2016).

279. Hagemans, M. L. C. *et al.* PAS-positive lymphocyte vacuoles can be used as diagnostic screening test for Pompe disease. *J. Inherit. Metab. Dis.* **33**, 133–139 (2010).

280. Schoser, B. G. H. *et al.* Adult-onset glycogen storage disease type 2: Clinico-pathological phenotype revisited. *Neuropathol. Appl. Neurobiol.* **33**, 544–559 (2007).

281. Musumeci, O. & Toscano, A. Diagnostic tools in late onset Pompe disease (LOPD). *Ann. Transl. Med.* **7**, 286–286 (2019).

282. Van den Hout, H. M. P. *et al.* The natural course of infantile Pompe's disease: 20 Original cases compared with 133 cases from the literature. *Pediatrics* **112**, 332–340 (2003).

283. Winkel, L. P. F. *et al.* The natural course of non-classic Pompe's disease; a review of 225 published cases. *Journal of Neurology* vol. 252 875–884 (2005).

284. Kishnani, P. S. *et al.* Pompe disease diagnosis and management guideline. *Genetics in Medicine* vol. 8 267–288 (2006).

285. Johnson, E. M. *et al.* Pulmonary function tests (maximum inspiratory pressure, maximum expiratory pressure, vital capacity, forced vital capacity) predict ventilator use in late-onset Pompe disease. *Neuromuscul. Disord.* **26**, 136–145 (2016).

286. Chan, J. *et al.* The emerging phenotype of late-onset Pompe disease: A systematic literature review. *Molecular Genetics and Metabolism* vol. 120 163–172 (2017).

287. Van Hout, H. Den *et al.* Recombinant human α-glucosidase from rabbit milk in Pompe patients. *Lancet* vol. 356 397–398 (2000).

288. van der Ploeg, A. T. *et al.* A Randomized Study of Alglucosidase Alfa in Late-Onset Pompe's Disease. *N. Engl. J. Med.* **362**, 1396–1406 (2010).

289. Güngör, D. *et al.* Impact of enzyme replacement therapy on survival in adults with Pompe disease: Results from a prospective international observational study. *Orphanet J. Rare Dis.* **8**, (2013).

290. De Vries, J. M. *et al.* Effect of enzyme therapy and prognostic factors in 69 adults with Pompe disease: An open-label single-center study. *Orphanet J. Rare Dis.* **7**, (2012).

291. Kishnani, P. S. *et al.* Cross-reactive immunologic material status affects treatment outcomes in Pompe disease infants. *Mol. Genet. Metab.* **99**, 26–33 (2010).

292. Pang, S. Y. Y. *et al.* The role of gene variants in the pathogenesis of neurodegenerative disorders as revealed by next generation sequencing studies: A review. *Transl. Neurodegener.* **6**, 1–11 (2017).

293. Liu, Y. T., Lee, Y. C. & Soong, B. W. What we have learned from the next-generation sequencing: Contributions to the genetic diagnoses and understanding of pathomechanisms of neurodegenerative diseases. *J. Neurogenet.* **29**, 103–112 (2015).

294. Lvovs, D., Favorova, O. O. & Favorov, A. V. A Polygenic Approach to the Study of Polygenic Diseases. *Acta Naturae* **4**, 59–71 (2012).

295. Timpson, N. J., Greenwood, C. M. T., Soranzo, N., Lawson, D. J. & Richards, J. B. Genetic architecture: The shape of the genetic contribution to human traits and disease. *Nat. Rev. Genet.* **19**, 110–124 (2018).

296. Chasioti, D., Yan, J., Nho, K. & Saykin, A. J. Progress in Polygenic Composite Scores in Alzheimer's and Other Complex Diseases. *Trends Genet.* **35**, 371–382 (2019).

297. Rahit, K. M. T. H. & Tarailo-Graovac, M. Genetic Modifiers and Rare Mendelian Disease.

Genes (Basel). **11**, 1–18 (2020).

298. Boycott, K. M. *et al.* International Cooperation to Enable the Diagnosis of All Rare Genetic Diseases. *American Journal of Human Genetics* vol. 100 695–705 (2017).

299. Chong, J. X. *et al.* The Genetic Basis of Mendelian Phenotypes: Discoveries, Challenges, and Opportunities. *American Journal of Human Genetics* vol. 97 199–215 (2015).

300. Raczy, C. *et al.* Isaac: Ultra-fast whole-genome secondary analysis on Illumina sequencing platforms. *Bioinformatics* **29**, 2041–2043 (2013).

301. McCann, E. P. *et al.* Evidence for polygenic and oligogenic basis of Australian sporadic amyotrophic lateral sclerosis. *J. Med. Genet.* (2020) doi:10.1136/jmedgenet-2020-106866.

302. Schmittgen, T. D. & Livak, K. J. Analyzing real-time PCR data by the comparative C(T) method. *Nat. Protoc.* **3**, 1101–8 (2008).

303. Wallner, B. & Elofsson, A. Pcons5: Combining consensus, structural evaluation and fold recognition scores. *Bioinformatics* **21**, 4248–4254 (2005).

304. Pires, D. E. V., Ascher, D. B. & Blundell, T. L. DUET: A server for predicting effects of mutations on protein stability using an integrated computational approach. *Nucleic Acids Res.* **42**, W314 (2014).

305. Pires, D. E. V., Ascher, D. B. & Blundell, T. L. MCSM: Predicting the effects of mutations in proteins using graph-based signatures. *Bioinformatics* **30**, 335–342 (2014).

306. Pandurangan, A. P., Ochoa-Montaño, B., Ascher, D. B. & Blundell, T. L. SDM: A server for predicting effects of mutations on protein stability. *Nucleic Acids Res.* **45**, W229–W235 (2017).

307. McDowall, J. & Hunter, S. InterPro protein classification. *Methods Mol. Biol.* **694**, 37–47 (2011).

308. Hancock, W. O. & Howard, J. Processivity of the motor protein kinesin requires two heads. *J. Cell Biol.* **140**, 1395–1405 (1998).

309. Miki, H., Okada, Y. & Hirokawa, N. Analysis of the kinesin superfamily: Insights into structure and function. *Trends in Cell Biology* vol. 15 467–476 (2005).

310. Filosto, M. *et al.* A Novel Mutation in the Stalk Domain of KIF5A Causes a Slowly Progressive Atypical Motor Syndrome. *J. Clin. Med.* **8**, 17 (2018).

311. Rinaldi, F. *et al.* A novel mutation in motor domain of KIF5A associated with an HSP/axonal neuropathy phenotype. *Journal of Clinical Neuromuscular Disease* vol. 16 153–158 (2015).

312. Lo Giudice, M. *et al.* A missense mutation in the coiled-coil domain of the KIF5A gene and late-onset hereditary spastic paraplegia. *Arch. Neurol.* **63**, 284–287 (2006).

313. Ebbing, B. *et al.* Effect of spastic paraplegia mutations in KIF5A kinesin on transport activity. *Hum. Mol. Genet.* **17**, 1245–1252 (2008).

314. Biton, A. *et al.* Polygenic Architecture of Human Neuroanatomical Diversity. *Cereb. Cortex* **30**, 2307–2320 (2020).

315. Sherry, S. T. *et al.* DbSNP: The NCBI database of genetic variation. *Nucleic Acids Res.* **29**, 308–311 (2001).

316. Capriotti, E., Fariselli, P., Rossi, I. & Casadio, R. A three-state prediction of single point mutations on protein stability changes. *BMC Bioinformatics* **9**, (2008).

317. De Filippi, P. *et al.* The angiotensin-converting enzyme insertion/deletion polymorphism modifies the clinical outcome in patients with Pompe disease. *Genet. Med.* **12**, 206–211 (2010).

318. Turner, S. M. F., Falk, D. J., Byrne, B. J. & Fuller, D. D. Transcriptome assessment of the Pompe (Gaa-/-) mouse spinal cord indicates widespread neuropathology. *Physiol. Genomics* **48**, 785–794 (2016).

319. Hesselink, R. P., Wagenmakers, A. J. M., Drost, M. R. & Van Der Vusse, G. J. Lysosomal

dysfunction in muscle with special reference to glycogen storage disease type II. *Biochimica et Biophysica Acta - Molecular Basis of Disease* vol. 1637 164–170 (2003).

320. Meinke, P., Limmer, S., Hintze, S. & Schoser, B. Assessing metabolic profiles in human myoblasts from patients with late-onset Pompe disease. *Ann. Transl. Med.* **7**, 277–277 (2019).

321. Grassano, M. *et al.* Mutational Analysis of Known ALS Genes in an Italian Population-Based Cohort. *Neurology* **96**, e600–e609 (2021).

322. Hou, L. *et al.* Screening of SOD1, FUS and TARDBP genes in patients with amyotrophic lateral sclerosis in central-southern China. *Sci. Rep.* **6**, (2016).

323. Kuuluvainen, L. *et al.* Oligogenic basis of sporadic ALS: The example of SOD1 p.Ala90Val mutation. *Neurol. Genet.* **5**, 335 (2019).

324. Yamashita, S. & Ando, Y. Genotype-phenotype relationship in hereditary amyotrophic lateral sclerosis. *Transl. Neurodegener.* **4**, (2015).

325. Del Bo, R. *et al.* TARDBP (TDP-43) sequence analysis in patients with familial and sporadic ALS: Identification of two novel mutations. *Eur. J. Neurol.* **16**, 727–732 (2009).

326. Corrado, L. *et al.* Mutations of FUS gene in sporadic amyotrophic lateral sclerosis. *J. Med. Genet.* **47**, 190–194 (2010).

327. Hewitt, C. *et al.* Novel FUS/TLS mutations and pathology in familial and sporadic amyotrophic lateral sclerosis. *Arch. Neurol.* **67**, 455–461 (2010).

328. Andersen, P. M. *et al.* Sixteen novel mutations in the Cu/Zn superoxide dismutase gene in amyotrophic lateral sclerosis: A decade of discoveries, defects and disputes. *Amyotrophic Lateral Sclerosis and Other Motor Neuron Disorders* vol. 4 62–73 (2003).

329. Millecamps, S. *et al.* SOD1, ANG, VAPB, TARDBP, and FUS mutations in familial amyotrophic lateral sclerosis: Genotype-phenotype correlations. *J. Med. Genet.* **47**, 554–560 (2010).

330. Canosa, A. *et al.* A familial ALS case carrying a novel p.G147C SOD1 heterozygous missense mutation with non-executive cognitive impairment. *Journal of Neurology, Neurosurgery and Psychiatry* vol. 85 1437–1439 (2014).

331. Origone, P. *et al.* Fast course ALS presenting with vocal cord paralysis: Clinical features, bioinformatic and modelling analysis of the novel SOD1 Gly147Ser mutation. *Amyotroph. Lateral Scler.* **13**, 144–148 (2012).

332. Osmanovic, A. *et al.* SPG7 mutations in amyotrophic lateral sclerosis: a genetic link to hereditary spastic paraplegia. *J. Neurol.* (2020) doi:10.1007/s00415-020-09861-w.

333. Panza, E., Martinelli, D., Magini, P., Vici, C. D. & Seri, M. Hereditary spastic paraplegia is a common phenotypic finding in ARG1 deficiency, P5CS deficiency and HHH syndrome: Three inborn errors of metabolism caused by alteration of an interconnected pathway of glutamate and urea cycle metabolism. *Front. Neurol.* **10**, 1–8 (2019).

334. Koh, K. *et al.* Novel mutations in the ALDH18A1 gene in complicated hereditary spastic paraplegia with cerebellar ataxia and cognitive impairment. *J. Hum. Genet.* **63**, 1009–1013 (2018).

335. Durand, C. M. *et al.* CYP2U1 activity is altered by missense mutations in hereditary spastic paraplegia 56. *Hum. Mutat.* **39**, 140–151 (2018).

336. Lek, M. *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291 (2016).

337. Stevanin, G. *et al.* Mutations in SPG11 are frequent in autosomal recessive spastic paraplegia with thin corpus callosum, cognitive decline and lower motor neuron degeneration. *Brain* **131**, 772–784 (2008).

338. Riboldi G. M. ; Di Fonzo A. B. GBA, Gaucher Disease, and Parkinson's Disease: From Genetic to Clinic to New Therapeutic Approaches. *Cells* **8**, 364 (2019).

339. Sato, C. *et al.* Analysis of the glucocerebrosidase gene in Parkinson's disease. *Mov. Disord.*

20, 367–370 (2005).

340. Cookson, M. R. *et al.* RING finger 1 mutations in Parkin produce altered localization of the protein. *Hum. Mol. Genet.* **12**, 2957–2965 (2003).

341. Manno, C., Lipari, A., Bono, V., Taiello, A. C. & La Bella, V. Sporadic Parkinson disease and Amyotrophic Lateral Sclerosis complex (Brait-Fahn-Schwartz Disease). *J. Neurol. Sci.* **326**, 104–106 (2013).

342. Calvo, A. *et al.* Parkinsonian traits in amyotrophic lateral sclerosis (ALS): a prospective population-based study. *J. Neurol.* **266**, 1633–1642 (2019).

343. Tan, E. K. *et al.* Analysis of PINK1 in Asian patients with familial parkinsonism [1]. *Clinical Genetics* vol. 68 468–470 (2005).

344. Marongiu, R. *et al.* PINK1 heterozygous rare variants: Prevalence, significance and phenotypic spectrum. *Hum. Mutat.* **29**, 565 (2008).

345. Newton, J., Milstien, S. & Spiegel, S. Niemann-Pick type C disease: The atypical sphingolipidosis. *Adv. Biol. Regul.* **70**, 82–88 (2018).

346. Hayashi, M. & Suzuki, T. Dyschromatosis symmetrica hereditaria. *J. Dermatol.* **40**, 336–343 (2013).

347. Kumar, K. R. *et al.* Defining the genetic basis of early onset hereditary spastic paraplegia using whole genome sequencing. *Neurogenetics* **17**, 265–270 (2016).

348. Drory, V. E., Birnbaum, M., Peleg, L., Goldman, B. & Korczyn, A. D. Hexosaminidase a deficiency is an uncommon cause of a syndrome mimicking amyotrophic lateral sclerosis. *Muscle and Nerve* **28**, 109–112 (2003).

349. Neudorfer, O. *et al.* Late-onset Tay-Sachs disease: Phenotypic characterization and genotypic correlation in 21 affected patients. *Genet. Med.* **7**, 119–123 (2005).

350. Figlewicz, D. A. & Orrell, R. W. The genetics of motor neuron diseases. *Amyotroph. Lateral Scler. Other Mot. Neuron Disord.* **4**, 225–231 (2003).

351. Valdmanis, P. N. & Rouleau, G. A. Genetics of familial amyotrophic lateral sclerosis. *Neurology* **70**, 144 LP – 152 (2008).

352. Nguyen, H. P., Van Broeckhoven, C. & van der Zee, J. ALS Genes in the Genomic Era and their Implications for FTD. *Trends in Genetics* vol. 34 404–423 (2018).

353. Van Blitterswijk, M. *et al.* Evidence for an oligogenic basis of amyotrophic lateral sclerosis. *Hum. Mol. Genet.* **21**, 3776–3784 (2012).

354. Huie, M. L., Shanske, A. L., Kasper, J. S., Marion, R. W. & Hirschhorn, R. A large Alu-mediated deletion, identified by PCR, as the molecular basis for glycogen storage disease type II (GSDII). *Hum. Genet.* **104**, 94–98 (1999).

355. Amiñoso, C. *et al.* Identification of the first deletion–insertion involving the complete structure of GAA gene and part of CCDC40 gene mediated by an Alu element. *Gene* **519**, 169–172 (2013).

356. Hoefsloot, L. H., Hoogeveen-Westerveld, M., Reuser, A. J. & Oostra, B. A. Characterization of the human lysosomal alpha-glucosidase gene. *Biochem. J.* **272**, 493–7 (1990).

357. Mechtler, T. P. *et al.* Neonatal screening for lysosomal storage disorders: feasibility and incidence from a nationwide study in Austria. *Lancet* **379**, 335–341 (2012).

358. Savarese, M. *et al.* Targeted gene panel screening is an effective tool to identify undiagnosed late onset Pompe disease. *Neuromuscul. Disord.* **28**, 586–591 (2018).

359. van der Ploeg, A. T. & Reuser, A. J. Pompe's disease. *The Lancet* vol. 372 1342–1353 (2008).

360. Fukuda, T. *et al.* Autophagy and Mistargeting of Therapeutic Enzyme in Skeletal Muscle in Pompe Disease. *Mol. Ther.* **14**, 831–839 (2006).

361. Raben, N. *et al.* Replacing acid α-glucosidase in Pompe disease: Recombinant and transgenic

enzymes are equipotent, but neither completely clears glycogen from type II muscle fibers. *Mol. Ther.* **11**, 48–56 (2005).

362. Fukuda, T. *et al.* Dysfunction of endocytic and autophagic pathways in a lysosomal storage disease. *Ann. Neurol.* **59**, 700–708 (2006).

363. Fukuda, T. *et al.* Autophagy and lysosomes in Pompe disease. *Autophagy* **2**, 318–320 (2006).

Scientific production arisen from this thesis

Peer-reviewed publications

1) Filosto M, Piccinelli SC, **Palmieri I**, *et al. A Novel Mutation in the Stalk Domain of KIF5A Causes a Slowly Progressive Atypical Motor Syndrome*. J Clin Med. 2018 Dec 22;8(1):17. doi: 10.3390/jcm8010017.

Abstracts at international meetings

1) **Palmieri I**, Conti S, Valente M, Pansarasa O, Diamanti L, Gagliardi S, Filosto M, Cereda C. *The genetic burden in an Italian cohort affected by Amyotrophic Lateral Sclerosis* 31st International Symposium on ALS/MND (2020) – Virtual
2) **Palmieri I**, Conti S, Pansarasa O, Diamanti L, Gagliardi S, Filosto M, Cereda C. *The genetic architecture and the role of comorbidity in the Amyotrophic Lateral Sclerosis.* XXIII Congresso della SIGU (2020) - Virtual
3) **Palmieri I**, Conti S, Pansarasa O, Diamanti L, Gagliardi S, Filosto M, Cereda C *The genetic burden in the amyotrophic lateral sclerosis clinical spectra of an Italian cohort* First annual meeting della Rete IRCCS delle neuroscienze e della neuroriabilitazione (2020) – Virtual
4) **Palmieri I**, Conti S, Pansarasa O, Diamanti L, Gagliardi S, Filosto M, Cereda C *The genetic architecture of amyotrophic lateral sclerosisand the role of comorbidity in Italian population* European Human Genetics Conference (ESHG 20202) – Virtual
5) Valente M, Zucca S, **Palmieri I**, Garau J, Rey F, Gagliardi S, Diamanti L, Ceroni M, Cereda C *Genetic testing of sporadic ALS patients reveals pathogenetic mutations in non-ALS genes* 28th International Symposium on ALS/MND (2017) - Boston, USA