U.PORTO

DOUTORAMENTO

CIÊNCIAS BIOMÉDICAS

Unravelling new spastic paraplegia genes and their functions through next generation sequencing

Sara Peres Morais

D

2018

U.PORTO

Sara Peres Morais. Unravelling new spastic paraplegia genes and their functions through next generation sequencing

Unravelling new spastic paraplegia genes and their functions through next generation sequencing

Sara Peres Morais

INSTITUTO DE CIÊNCIAS BIOMÉDICAS ABEL SALAZAR

D.ICBAS **2018**

U.PORTO

INSTITUTO DE CIÊNCIAS BIOMÉDICAS ABEL SALAZAR
UNIVERSIDADE DO PORTO

Unravelling new spastic paraplegia genes and their functions through next generation sequencing

Sara Peres de Morais

Tese de Doutoramento em Ciências Biomédicas
Instituto de Ciências Biomédicas Abel Salazar

2018

SARA PERES DE MORAIS

UNRAVELLING NEW SPASTIC PARAPLEGIA GENES AND THEIR FUNCTIONS THROUGH NEXT GENERATION SEQUENCING

Tese de Candidatura ao grau de Doutor em Ciências Biomédicas, submetida ao Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto.

Orientadora – Doutora Isabel da Conceição Moreira Pereira Alonso

Investigadora do Instituto de Biologia Molecular e Celular e Professora Afiliada no Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto.

Coorientador – Doutor Giovanni Stevanin

Diretor de Pesquisas do INSERM e EPHE

Coorientador - Professor Doutor António Jorge dos Santos Pereira de Sequeiros

Professor Catedrático do Instituto de Ciências Biomédicas de Abel Salazar da Universidade do Porto.

“We do not need magic to change the world, we carry all the power we need inside ourselves already: we have the power to imagine better.”

J K Rowling

Table of Contents

Publications i
Abbreviations iii
Abstract ix
Resumo x
1. Introduction 1
1.1. Hereditary spastic paraplegias 2
1.1.1. Definition and prevalence 2
1.1.2. Clinical aspects 2
1.1.3. Genetic aspects 4
1.2. Clinical diagnosis and heterogeneity 9
1.2.1 Clinical diagnosis and HSPs differential diagnostics 9
1.2.2 Clinical, genetic and molecular heterogeneity 10
1.3. Mechanisms of pathogenesis in the HSPs 11
1.3.1 Axonal transport and microtubule dynamics 12
1.3.2 Endoplasmic reticulum morphogenesis 14
1.3.3 Membrane trafficking 15
1.3.4 Mitochondrial function 18
1.3.5 Lipid metabolism 19
1.3.6 Myelination and development 21
1.3.7 Nucleotide metabolism and DNA repair 22
1.3.8 Autophagy 22
1.4. Animal models 24
1.4.1 HSP animal models 24
1.4.2 Zebrafish model 25
2. Major goals 1
2.1 Main aims 2
2.2 Specific objectives 2
3. Material and Methods 3
3.1. Patients 4
3.1.1 Inclusion criteria 4
3.1.2 Clinical phenotypes 4
3.1.3 Ethics 5
3.2. Biological material 5
3.2.1 DNA extraction and quality check 5
3.2.2 Skin biopsies and fibroblast culture 5
3.3. Targeted next generation sequencing for screening of known HSP genes 6
3.3.1 Panel design 6

3.3.2 Experimental procedure 6
3.3.3 Bioinformatic analysis..... 7
3.3.4 Confirmation and segregation analysis..... 8
3.4. Whole exome sequencing (WES) for identification of new HSP genes..... 9
3.4.1 Family selection 9
3.4.2 Experimental procedure 9
3.4.3 Bioinformatic analysis.....10
3.4.4 Confirmation and segregation analysis.....10
3.5. Genome-wide genotyping and linkage analysis.....11
3.5.1 Family selection 11
3.5.2 Experimental procedure 11
3.5.3 Bioinformatic analysis.....11
3.6. Functional studies 12
3.6.1 DNA constructs and site-directed mutagenesis 12
3.6.2 Cell culture and transfection (HEK293T and fibroblasts) 12
3.6.3 Antibodies and subcellular fractionation 12
3.6.4 Western blot.....13
3.6.5 Immunofluorescence.....13
3.6.6 RNA extraction, RT-PCR and qPCR 14
3.6.7 Zebrafish studies.....14
3.7. Targeted next generation sequencing for screening of new candidate HSP genes (Fluidigm assay) 15
3.7.1 Panel design and family selection 15
3.7.2 Experimental protocol.....15
3.7.3 Bioinformatic analysis.....16
3.7.4 Confirmation and segregation analysis and *AIMP1* study.....16
4. Results.....17
4.1 Search for molecular diagnosis in already identified genes 18
4.2 Search for new HSP causing genes.....35
4.2.1 *ALDH18A1* as a causative gene for HSPs.....37
4.2.2 *AIMP1* as a causative gene for HSPs.....57
4.2.3 *AKAP6* as a candidate for HSPs 67
4.2.4. Results from the remaining 9 families studied by whole exome sequencing....79
4.3 Search for additional families with causative variants in the new candidate genes .89
4.3.1 Mini panel results 90
5. Discussion.....97
5.1 Genetic characterization of the Portuguese cohort with HSP 98
5.1.1 Overview of the molecular strategies and results.....98
5.1.2 Gene frequencies in the Portuguese cohort.....101
5.1.3 Molecular heterogeneity 103

5.2 Molecular approaches for HSP diagnosis 104
5.2.1 Unravelling “old” and new spastic paraplegia genes 104
5.2.2 Next generation sequencing challenges 107
5.3 Unravelling spastic paraplegia mechanisms 110
5.3.1 Pathophysiological mechanisms involved in HSP and axonal degeneration 110
5.3.2 New spastic paraplegia genes and mechanisms 111
6. Conclusions 115
6.1 Final remarks 116
6.2 Future perspectives 117
References 116
Annex 146
Tables 147
Supplementary Articles 150
Annex 2 - Genomic mechanisms underlying PARK2 large deletions identified in a cohort of patients with PD 150
Annex 3 - Large-Scale Functional RNAi screen in *C. elegans* Identifies TGF-ß and Notch Signaling Pathways as Modifiers of CACNA1A 157
Annex 4 – Chapter 25.3 Hereditary Spastic Paraplegias 167
Acknowledgments 173

Publications

In this thesis were used results already published or in preparation as described below:

Article 1: Morais S, Raymond L, Mairey M, Coutinho P, Brandão E, Ribeiro P, Loureiro JL, Sequeiros J, Brice A, Alonso I, Stevanin G: "Massive Sequencing of 70 Genes Reveals a Myriad of Missing Genes or Mechanisms to be uncovered in Hereditary Spastic Paraplegia". European Journal of Human Genetics 2017 Nov;25(11):1217-1228. doi: 10.1038/ejhg.2017.124

Article 2: Coutelier M, Goizet C, Durr A, Habarou F, Morais S, Dionne-Laporte A, Tao F, Konop J, Stoll M, Charles P, Jacoupy M, Matusiak R, Alonso I, Tallaksen C, Mairey M, Kennerson M, Gaussen M, Schule R, Janin M, Morice-Picard F, Durand C, Depienne C, Calvas P, Coutinho P, Saudubray JM, Rouleau G, Brice A, Nicholson G, Darios F, Loureiro JL, Zuchner S, Ottolenghi C, Mochel F, Stevanin G: "Alteration of ornithine metabolism leads to dominant and recessive hereditary spastic paraplegia". Brain 2015 Aug;138(Pt 8):2191-205. doi: 10.1093/brain/aww143

Article 3: Morais S, Tesson C, Coutinho P, Brandão E, Ribeiro P, Loureiro JL, Sequeiros J, Brice A, Stevanin G, Alonso I. "AIMP1, a gene implicated in a range of phenotypes from intellectual disability to Hereditary Spastic Paraplegia". Submitted

In Annex we show the papers submitted during the PhD on works performed previously (annexes 2 and 3) and a book chapter co-authored during the PhD (annex 4):

Annex 2: Morais S, Bastos-Ferreira R, Sequeiros J, Alonso I: "Genomic mechanisms underlying *PARK2* large deletions identified in a cohort of patients with PD". Neurol Genet. May 3; 2(3): e73. doi: 10.1212/NXG.0000000000000073

Annex 3: Pereira Mda C, Morais S, Sequeiros J, Alonso I: "Large-Scale Functional RNAi screen in *C. elegans* Identifies TGF-ß and Notch Signaling Pathways as Modifiers of CACNA1A". ASN Neuro. 2016 Mar 22;8(2). doi: 10.1177/1759091416637025

Annex 4: Matilla-Dueñas A, Corral-Juan M, Seuma A, Rolán D, Ispierto L, Morais S, Sequeiros J, Alonso I, Bertrán V, Serrano-Munuera C, Pintos-Morell G, Álvarez R, Sánchez I. “Chapter 25: Rare Neurodegenerative Diseases: Clinical and Genetic Update”. Book: Rare Diseases Epidemiology: Update and Overview, Publisher: Springer, Editors: Manuel Posada, Domenica Taruscio, Stephen Groft. DOI: 10.1007/978-3-319-67144-

Abbreviations

ABCA12	ATP Binding Cassette Subfamily A Member 12
ABCD1	ATP Binding Cassette Subfamily D Member 1
AC5	Adenylyl cyclase type 5
AD	Autosomal Dominant
ADAR	Adenosine Deaminase, RNA Specific
AFG3L2	AFG3 like Matrix AAA Peptidase Subunit 2
AIMP1	Aminoacyl TRNA Synthetase Complex Interacting Multifunctional Protein 1
AKAP6	A-Kinase Anchoring Protein 6
ALDH18A1	Aldehyde Dehydrogenase 18 Family Member A1
ALS	Amyotrophic Lateral Sclerosis
ALS2	Amyotrophic Lateral Sclerosis 2
AMPD2	Adenosine Monophosphate Deaminase 2
AO	Age at Onset
AP	Adaptor Protein Complex
AP4B1	Adaptor Related Protein Complex 4 Beta 1 Subunit
AP4E1	Adaptor Related Protein Complex 4 Epsilon 1 Subunit
AP4M1	Adaptor Related Protein Complex 4 Mu 1 Subunit
AP4S1	Adaptor Related Protein Complex 4 Sigma 1 Subunit
AP5B1	Adaptor Related Protein Complex 5 Beta 1 Subunit
AP5M1	Adaptor Related Protein Complex 5 Mu 1 Subunit
AP5S1	Adaptor Related Protein Complex 5 Sigma 1 Subunit
AP5Z1	Adaptor Related Protein Complex 5 Zeta 1 Subunit
AR	Autosomal Recessive
ARG	Arginine
ARG1	Arginase 1
ARL6IP1	ADP Ribosylation Factor Like GTPase 6 Interacting Protein 1
ARSACS	AR spastic ataxia of Charlevoix-Saguenay
ARSI	Arylsulfatase Family Member I
ATL1	Atlastin GTPase 1
ATP	Adenosine Triphosphate
ATP13A2	ATPase 13A2
B4GALNT1	Beta-1,4-N-Acetyl-Galactosaminyltransferase 1
BICD2	Bicaudal D Cargo Adaptor 2
BMP	Bone Morphogenic Protein

BMP6	Bone Morphogenic Protein 6
BSA	Bovine Serum Albumin
BSCL2	Berardinelli-Seip Congenital Lipodystrophy 2
C12orf65	Chromosome 12 open reading frame 65
C19orf12	Chromosome 19 open reading frame 12
CAPN1	Calpain 1
CC	Corpus callosum
CCFS	Composite cerebellar function severity score
CCT5	Chaperonin Containing TCP1 Subunit 5
cDNA	complementary DNA
Chr	Chromosome
CITRU	Citrulline
CLN8	Ceroid-Lipofuscinosis, Neuronal 8
CMT	Charcot-Marie-Tooth
CNS	Central Nervous System
COPI	Coat protein I
CPS1	Carbamoylphosphate synthase 1
CPT1C	Carnitine Palmitoyltransferase 1C
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CYP2U1	Cytochrome P450 Family 2 Subfamily U Member 1
CYP7B1	Cytochrome P450 Family 7 Subfamily B Member 1
DD	Developmental delay
DDHD1	DDHD Domain Containing 1
DDHD2	DDHD Domain Containing 2
DHEA	Dehydroepiandrosterone
DMEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic Acid
DNM2	Dynamin 2
dNTP	Deoxynucleotide triphosphate
DYT	Dystonia
EIF2B5	Eukaryotic Translation Initiation Factor 2B Subunit Epsilon
ELOVL4	Elongation of Very Long Chain Fatty Acids Protein 4
ENMG	Electro-neuro-myogram
ENTPD1	Ectonucleoside Triphosphate Diphosphohydrolase 1
ER	Endoplasmic Reticulum
ERAD	ER-associated degradation
ERLIN1	ER Lipid Raft Associated 1

ERLIN2	ER Lipid Raft Associated 2
ESCRT	Endosomal sorting complexes required for transport
ExAC	Exome Aggregation Consortium
EXOSC3	Exosome Component 3
FA2H	Fatty Acid 2-Hydroxylase
FAM134B	Family with Sequence Similarity 134 Member B
FARS2	Phenylalanyl-TRNA Synthetase 2, Mitochondrial
FBXO7	F-Box Protein 7
FLRT1	Fibronectin Leucine Rich Transmembrane Protein 1
FTD	Frontotemporal Dementia
GAD1	Glutamate Decarboxylase 1
GBA2	Glucosylceramidase Beta 2
GCH1	GTP Cyclohydrolase 1
GJA1	Gap Junction Protein Alpha 1
GJC2	Gap Junction Protein Gamma 2
GLB1	Galactosidase Beta 1
gnomAD	Genome Aggregation Database
GRID2	Glutamate Ionotropic Receptor Delta Type Subunit 2
HC	Head circumference
Het	Heterozygous
HIST1H2AM	Histone Cluster 1 H2A Family Member M
HIV	Human immunodeficiency virus
HLD	Hypomyelinating Leukodystrophy
Hom	Homozygous
HSN	Hereditary Sensory Neuropathy
HSP	Hereditary Spastic Paraplegia
HSPD1	Heat shock 60 kDa protein 1
HTLV-I	Human T-cell leukemia virus 1
IBA57	IBA57 Homolog, Iron-Sulfur Cluster Assembly
ID	Intellectual Disability
IFIH1	Interferon Induced with Helicase C Domain 1
IFT74	Intraflagellar Transport 74
JMJD4	Jumonji Domain Containing 4
KANK1	KN Motif and Ankyrin Repeat Domains 1
KCNA2	Potassium Voltage-Gated Channel Subfamily A Member 2
KIAA0196	Strumpellin gene
KIDINS220	Kinase D Interacting Substrate 220

KIF1A	Kinesin Family Member 1A
KIF1C	Kinesin Family Member 1C
KIF5A	Kinesin Family Member 5A
KLC4	Kinesin Light Chain 4
L	Left
L1CAM	L1 Cell Adhesion Molecule
LDs	Lipid droplets
LL	Lower Limbs
LOD	Logarithm of the odds
LYST	Lysosomal Trafficking Regulator
MAF	Minor Allele Frequency
MAG	Myelin Associated Glycoprotein
MAP1A	Microtubule Associated Protein 1A
MARS	Methionyl-TRNA Synthetase
MBD	Microtubule binding domain
MD	Motor delay
MIT	Microtubule interacting and endosomal trafficking domain
Mito	Mitochondrial
MLPA	Multiplex ligation-dependent probe amplification
MO	Morpholino
MRC	Medical Research Council
MRI	Magnetic Resonance Imaging
mRNA	messenger Ribonucleic acid
MT	Mutant
MT-ATP6	Mitochondrially Encoded ATP Synthase 6
MT-CO3	Mitochondrially Encoded Cytochrome C Oxidase III
mtDNA	mitochondrial DNA
MTHFR	Methylenetetrahydrofolate Reductase
MT-TE	Mitochondrially Encoded TRNA Glutamic Acid
MT-TI	Mitochondrially Encoded TRNA Isoleucine
N	Number
NA	Not Available
NAGS	N-acetylglutamate synthetase
NBIA	Neurodegeneration with Brain Iron Accumulation
ND	Not done
NGS	Next-generation sequencing
NIPA1	Non Imprinted in Prader–Willi/Angelman Syndrome 1

NPB	Neuropeptide B
NT5C2	5'-Nucleotidase, Cytosolic II
OPA3	Optic Atrophy Protein 3
ORN	Ornithine
OTC	Ornithine transcarbamylase
OXPHOS	Oxidative phosphorylation
P5C	Pyrroline-5-carboxylate
P5CS	Delta-1-pyrroline-5-carboxylate synthase
PBS-T	Phosphate buffered saline with 1% Tween-20
PCR	Polymerase Chain Reaction
PD	Parkinson´s disease
PDE4D3	Phosphodiesterase type 4D3
PDK1	3-phosphoinositide-dependent protein kinase-1
PGAP1	Post-GPI Attachment to Proteins 1
PKA	Protein Kinase A
PLA2G6	Phospholipase A2 Group VI
PLP1	Proteolipid Protein 1
PLS	Primary Lateral Sclerosis
PMD	Psychomotor delay
PMLD	Pelizaeus-Merzbacher-like Disease
PNPLA6	Patatin like Phospholipase Domain Containing 6
PNS	Peripheral nervous system
PRO	Proline
PYCR2	Pyrroline-5-Carboxylate Reductase Family Member 2
PVDF	Polyvinylidene fluoride
qPCR	quantitative real-time PCR
R	Right
RAB3GAP2	RAB3 GTPase Activating Non-Catalytic Protein Subunit 2
REEP1	Receptor Expression Enhancing Protein 1
REEP2	Receptor Expression Enhancing Protein 2
RNA	Ribonucleic acid
RNASEH2B	Ribonuclease H2 Subunit B
RT-PCR	Reverse transcriptase-PCR
RTN2	Reticulon 2
RyR2	Ryanodine receptor type 2
SACS	Sacsin Molecular Chaperone

SAMHD1	SAM and HD Domain Containing Deoxynucleoside Triphosphate Triphosphohydrolase 1
SCAs	Spinocerebellar Ataxias
SERINC4	Serine Incorporator 4
SETX	Senataxin
SGSM3	Small G Protein Signaling Modulator 3
SLC16A2	Solute carrier family 16, member 2
SLC33A1	Solute carrier family 33
SNP	Single Nucleotide Polymorphism
SPAST	Spastin gene
SPATAX	European and Mediterranean spastic paraplegia and ataxia network
SPG	Spastic Paraplegia Gene
SPG Un	Unassigned Spastic Paraplegia Gene
SPG7	Spastic Paraplegia 7
SPG11	Spastic Paraplegia 11
SPG20	Spastic Paraplegia 20
SPG21	Spastic Paraplegia 21
SUCO	SUN Domain Containing Ossification Factor
TCC	Thin corpus callosum
TECPR2	Tectonin Beta-Propeller Repeat Containing 2
TFG	TRK-Fused Gene
TUBB4A	Tubulin Beta 4A Class IVa
UL	Upper Limbs
UCHL1	Ubiquitin C-Terminal Hydrolase L1
USP8	Ubiquitin Specific Peptidase 8
VCP	Valosin-containing protein
VPS37A	Vacuolar Protein Sorting-Associated Protein 37A
VUS	Variants of Unknown Significance
WASHC5	WASH Complex Subunit 5
WES	Whole Exome Sequencing
WDR48	WD Repeat Domain 48
WM	White Matter
WT	Wild Type
X-L	X-Linked
ZFYVE26	Zinc finger, FYVE domain containing 26
ZFYVE27	Zinc finger, FYVE domain containing 27
ZFR	Zinc Finger RNA Binding Protein

Abstract

Hereditary spastic paraplegias (HSP) are a clinically and genetically heterogeneous group of rare neurodegenerative disorders characterized by progressive lower limb spasticity and weakness. The prevalence of HSP has been estimated at 3-10/100,000 in Europe and of 4.1/100 000 in Portugal. Genetically, HSPs are complex with all modes of inheritance described and with at least 79 SPG-associated *loci* and more than 70 identified genes. Although a high number of genes have been identified, still 62.9% of the Portuguese families with HSP remained without a molecular diagnosis after the genetic screening of the most frequently mutated genes.

Our aim was so to provide a molecular diagnosis to a cohort of 100 Portuguese HSP families without molecular diagnosis, not only by searching for variants in genes already associated with HSP but, additionally, by searching for new causative genes. We used next generation sequencing (panel and exome sequencing) combined with linkage analysis. Functional characterization was also performed for two of the novel genes identified (*AIMP1* and *AKAP6*). In this study, we were able to deliver a molecular diagnosis to 24 families, increasing to 52,2% the diagnosis rate in our HSP cohort and associating for the first time *AIMP1* variants with HSP, thus expanding the phenotypic spectrum for variants in this gene. Additional candidate genes were identified in 7 families, two of which are strong candidates (*AKAP6* and *SUCO*), that will be further studied in order to ascertain their pathogenic role. Several variants of unknown significance were also identified that could potentially increase the number of diagnosis after a more in-depth study, which also includes a prospective new mode of inheritance for *KIF5A* variants to be proven. This study further confirmed the high genetic heterogeneity present in HSPs, most specifically in the Portuguese cohort, proving that a considerable number of genes and variants are apparently orphan, affecting single families.

This is the most extended genetic screening in a large set of HSP families in Portugal. In addition to the molecular diagnosis identified, it describes *AIMP1* as a new HSP-causative gene and identifies several new potential disease-causing genes, improving the knowledge on the genetics of the HSP. The identification of new genes implicated in HSPs is fundamental to understand the cellular biology and pathogenesis of this group of disorders that remains largely elusive. So, in addition to the direct impact on the patients and families, this study can have a major impact in the HSP field since, it can help to decipher the molecular pathways leading to neurodegeneration in HSP and other neurodegenerative disorders.

Resumo

As paraparésias espásticas hereditárias (HSPs) são um grupo de doenças neurodegenerativas raras, clínica e geneticamente heterogéneo, caracterizadas por espasticidade e fraqueza progressiva dos membros inferiores. A prevalência das HSPs foi estimada em 3-10 / 100.000 na Europa e 4,1 / 100.000 em Portugal. Geneticamente, as HSPs são complexas, com todos os modos de hereditariedade descritos e com, pelo menos, 79 *loci* associados e mais de 70 genes identificados. Embora já tenha sido identificado um grande número de genes, 62,9% das famílias portuguesas com HSP ainda permaneciam sem um diagnóstico molecular após o rastreio genético dos genes mais frequentemente mutados.

O nosso objetivo foi estabelecer o diagnóstico molecular em 100 famílias portuguesas com HSP que permaneciam sem diagnóstico molecular, não só pela pesquisa de variantes em genes já associados às HSPs, mas também, pela pesquisa de novos genes responsáveis pela doença. Utilizamos sequenciação de nova geração (sequenciação de painéis e exomas) combinada com a análise de ligação. A caracterização funcional foi também realizada para dois dos novos genes identificados (*AIMP1* e *AKAP6*).

Com este estudo fomos capazes de estabelecer o diagnóstico molecular em 24 famílias, aumentando assim para 52,2% a taxa de diagnóstico na coorte Portuguesa de famílias com HSP e associamos pela primeira vez, variantes no gene *AIMP1* com HSP, expandindo assim o espectro fenotípico para variantes neste gene. Foram ainda identificados genes candidatos em 7 famílias, sendo dois deles fortes candidatos (*AKAP6* e *SUCO*) que continuarão a ser estudados com o objetivo de determinar o seu papel patogénico. Diversas variantes de significado desconhecido foram também identificadas que, após um estudo mais aprofundado, podem potencialmente resultar num aumento do número de diagnósticos alcançados e que também inclui um potencial novo modo de transmissão, a ser comprovado, para variantes presentes no gene *KIF5A*. Este estudo confirmou ainda a grande heterogeneidade genética presente nas HSPs, mais especificamente na coorte portuguesa, provando que um número considerável de genes e variantes são aparentemente órfãos, afetando famílias singulares.

Este estudo é o rastreio genético mais alargado num grande conjunto de famílias com HSP em Portugal. Para além dos diagnósticos moleculares alcançados, descreve o gene *AIMP1* como um novo gene responsável por HSP e identifica vários novos genes como potenciais causadores da doença, melhorando o conhecimento sobre a genética das HSPs. A identificação de novos genes implicados nas HSPs é fundamental para a compreensão da biologia celular e patogénese desse grupo de doenças que permanece em grande parte desconhecida. Assim, além do impacto direto nos doentes e nas suas famílias, este estudo pode ter também um grande impacto no universo das HSPs, uma vez que pode ajudar a

decifrar as vias moleculares que levam à neurodegeneração nesta doença assim como noutras doenças neurodegenerativas.

1. Introduction

1.1. Hereditary spastic paraplegias

1.1.1. Definition and prevalence

Hereditary spastic paraplegias (HSP), also known as hereditary spastic paraparesis, familial spastic paraplegia or Strümpell-Lorrain disease, is a group of rare neurodegenerative disorders that cause progressive weakness and spasticity of the lower limbs. The first description of the disease is considered to be of Strümpell in 1880 (Strümpell, 1880) although there is a report of a similar disorder by Seeligmuller in 1876 (Seeligmuller A., 1876). These first cases described by Strümpell and later by Lorrain (Lorrain, 1898) presented as only pure cases with a probable autosomal dominant inheritance. Later, complicated cases of spastic paraplegia appeared in the literature and a more vast definition of HSP was defined (Harding, 1981, 1983; Holmes and Shaywitz, 1977). HSP started to be clinically subdivided in pure or complex forms according to the absence or presence of additional neurological or extraneurological symptoms like ataxia, neuropathy, among others.

HSP are now clinically and genetically very heterogeneous with more than 70 genes identified, covering all modes of inheritance. The neuropathological hallmark of the disease is the degeneration of the long descending fibers of the corticospinal tract and posterior columns (Souza et al., 2016).

HSP is the second most frequent group of motor neuron disorders (Dion et al., 2009) with an estimated global prevalence of 4.26/100 000 ranging from 0.9-9.6/100 000 depending on the geographic area (Ruano et al., 2014) and is of 4.1/100 000 in Portugal (Coutinho et al., 2013).

1.1.2. Clinical aspects

The symptoms that characterize a pure or “uncomplicated” form of HSP are bilateral lower limb spasticity (maximal at the hamstring, quadriceps, adductor, and gastrocnemius-soleus muscles) more prominent when walking than at rest, and weakness (greatest in the iliopsoas, hamstring, and tibialis anterior muscles). Additionally, increased lower limb muscle tone, hyperreflexia, extensor plantar responses and lower-limb sensory disturbances (like attenuated vibratory sensation in the ankles and toes) are frequently found. The onset is subtle, and the disease is usually slowly progressive with patients complaining of leg stiffness, abnormal wear of the outside of the shoes, frequent falls, cramps, abnormal gait or gait instability. The progressive walking difficulty usually leads to the need of canes, walkers or wheelchairs. It is also frequent to find patients complaining of urinary urgency that could be an

early symptom of the disease and the presence of *pes cavus* is also a common finding (Fink, 2013; Parodi et al., 2017; Salinas et al., 2008; Souza et al., 2016).

In the complex or “complicated” forms there are additional neurological or systemic features such as cognitive impairment (intellectual disability or dementia), ataxia, neuropathy, epilepsy, optic atrophy, muscular atrophies, distal wasting, bradykinesia, parkinsonism and dystonia. With the use of ancillary tests such as brain and spinal cord magnetic resonance imaging (MRI), some neuroimaging abnormalities could be found in several patients such as a thinning of the *corpus callosum (TCC)* (commonly present in patients from some recessive HSP forms, mostly caused by *SPG11* and *SPG15* disease-causing variants), leukodystrophy, spinal cord atrophy, brain iron accumulation, hypomyelination, hydrocephalus and cerebellar atrophy (Souza et al., 2016). Extra-neurological manifestations could also occur and cause vision loss (retinitis pigmentosa, cataracts, optic neuropathy, optic atrophy and macular degeneration), skin changes, dysmorphic features and orthopedic abnormalities, although more uncommon.

Age at onset of HSP is quite variable, ranging from the first years of life to a late onset. Even when the genetic cause is known, there is a broad range of age at onset found, even among those with the same genetic background, and a high intrafamilial variability is frequently observed. Asymptomatic carriers are often found, suggesting the influence of still unidentified modifying factors. Still, autosomal recessive forms are most commonly associated with an early onset, which could not be used for genotype prediction as several dominant forms like SPG3A have an expected childhood onset and some recessive forms like SPG7 or SPG44 have an expected adulthood onset. Also, an earlier disease onset was associated with a less severe presentation (Schüle et al., 2016).

Progression in HSP is also very irregular with patients living with a relatively non-progressive disability for many years and others experimenting a rapid worsening. It is also common to find an inconstant progression, where patients have a fast worsening during adolescence and after several years a slower progression rate. Many patients seem to reach an apparent disease stabilization, which could be the result of multiple factors like a possible functional compensation from neuroplasticity (Fink, 2013). This wide spectrum of age at onset, progression and associated symptoms allied to a large number of genetic causes, in which several genes have only a few families identified, impairs the establishment of phenotype–genotype predictions. Additionally, as mutations in several genes have been identified as being responsible for both pure and complex phenotypes further complicates pointing out a possible genetic cause.

Presently, there is no specific treatment to prevent or slow the neuronal degeneration. The only strategy consists in symptomatic treatment like antispasticity drugs, botulinum toxin and orthoses. Physiotherapy and rehabilitation are also important to improve the patients quality of life (Fink, 2013; Parodi et al., 2017; Souza et al., 2016).

1.1.3. Genetic aspects

HSP are genetically very heterogeneous with additional new genes constantly being associated with the disease. Until now, there are 79 HSP *loci* with 65 identified genes. Additionally, at least 39 more genes were associated with symptoms mimicking part of the spastic paraplegia clinical spectrum either associated to HSP only (SPG Un, n=12) or already associated with several other diseases (MM_SPG, n=27). All of them are summarized in table 1. These genes can be inherited by a variety of ways that include all classical modes of transmission: autosomal dominant (AD), autosomal recessive (AR), X-linked, maternally inherited and possibly also by imprinting-like inheritance (*KANK1* gene) (Lerer et al., 2005). Some of the genes have already been described as having more than one inheritance mode such as *KIF1A*, for which it was found that some missense variants in the motor domain of the protein result in a dominant inheritance mode (Lee et al., 2015a) or *REEP2* and *ALDH18A1* genes (Coutelier et al., 2015; Esteves et al., 2014) that are associated to AD and AR inheritance modes regardless of the protein domain affected, among others.

The most common cause in the dominant forms are variants in the *SPAST* (SPG4) gene, responsible for 17 to 26% of all the HSP cases and causing a pure form of HSP. Following it, causative variants in the *ATL1* (SPG3A), *REEP1* (SPG31) and *KIF5A* (SPG10) genes are also common, being (together with SPG4) responsible for almost 57% of the families with a dominant form of HSP (Hensiek et al., 2015; Parodi et al., 2017). The frequency of AD inherited forms of *KIF1A* mutations are not yet precisely known but unpublished series form the SPATAX international network indicate that this form may be the second most common AD HSP. For the recessive forms, the most commonly mutated gene is *SPG11*, being responsible for around 21% of the cases and causing a complex form of the disease (Stevanin et al., 2008). The remaining most frequent causes that account for 25% of the AR-HSP families are the genes *CYP7B1* (SPG5), *SPG7* and *ZFYVE26* (SPG15) (Tesson et al., 2015) to which can be added mutations in *SACS* in spastic ataxias in several countries.

In HSP, similarly to the hereditary ataxias, a classification by their genetic loci was started, designated by SPG (from the spastic paraplegia initials) followed by a number (according to the chronological order of the discovery). The discovery of all the first HSP genes was possible through linkage analysis. This strategy allowed the identification of the genomic region where the causative genes are located, by genotyping genetic markers and studying their segregation on the affected families' pedigrees. Nowadays, with next-generation sequencing (NGS), more genes and candidate genes are constantly being identified, which led us to this extraordinary number of HSP disease-causing genes already known. These new approaches: panel sequencing, whole-exome and whole-genome sequencing allow the study of the less frequently mutated genes (only found in few, or even single families), and also the study of all

the already identified HSP-causing and associated genes (independently of the inheritance mode and clinical type of HSP) which could greatly improve the capacity to deliver a molecular diagnosis to the patients and their families. Still, nowadays, too many HSP families remain without genetic diagnosis. This could be due to several factors: (1) patients were only screened for the most frequently HSP causing genes by Sanger sequencing, (2) causative variants are present in a still not identified HSP-causing gene, (3) the causative variant is a genomic rearrangement (large deletion or duplication) lost by the limitations of the NGS techniques, (4) lack of proof of pathogenicity of the variants found, mostly regarding missense variants, being reported only as variants of uncertain significance (VUS) until further genetic or functional study, (5) genetic cause not exonic or even digenic/oligogenic inheritance, increasing the difficulty in the identification and in proving their pathogenicity.

Sporadic cases are by far the most frequent in clinical practice but they represent a real challenge. They were usually not studied due to the lack of family history and so, proof of a genetic origin, which helps to discard other differential diagnosis. However, the presence of *de novo* variants, compound heterozygosity, variable expressivity, early death of the transmitting parent and reduced or age-dependent penetrance are common findings that could be masking a familial disease (Souza et al., 2016; Tesson et al., 2015) and should be considered (mostly for panel sequencing strategies). Indeed, *SPAST* and *SPG11* variants are frequent in sporadic cases (Depienne et al., 2006; Stevanin et al., 2008), which is also true for *KIF1A* with numerous *de novo* cases reported (Cheon et al., 2017; Citterio et al., 2015).

Table 1: List of HSP associated genes

Type	Chr	Gene	Phenotype	Inheritance mode	Reference
SPG1	Xq28	***L1CAM***	Complex	X-linked	(Jouet et al., 1994)
SPG2	Xq22.2	***PLP1***	Pure or Complex	X-linked	(Saugier-Veber et al., 1994)
SPG3A	14q22.1	***ATL1***	Pure or Complex	Autosomal dominant or recessive	(Khan et al., 2014; Zhao et al., 2001)
SPG4	2p22.3	***SPAST***	Pure or Complex	Autosomal dominant	(Hazan et al., 1999)
SPG5A	8q21.3	***CYP7B1***	Pure or Complex	Autosomal recessive	(Tsaousidou et al., 2008)
SPG6	15q11.1	***NIPA1***	Pure or Complex	Autosomal dominant	(Rainier et al., 2003)
SPG7	16q24.3	***SPG7***	Pure or Complex	Autosomal recessive or dominant	(Casari et al., 1998; Sánchez-Ferrero et al., 2013)
SPG8	8q24.13	***WASHC5***	Pure	Autosomal dominant	(Valdmanis et al., 2007)
SPG9	10q24.1	***ALDH18A1***	Pure or Complex	Autosomal dominant or recessive	(Coutelier et al., 2015)
SPG10	12q13	***KIF5A***	Pure or Complex	Autosomal dominant	(Reid et al., 2002)
SPG11	15q21.1	***SPG11***	Pure or Complex	Autosomal recessive	(Stevanin et al., 2007a)
SPG12	19q13	***RTN2***	Pure	Autosomal dominant	(Montenegro et al., 2012)
SPG13	2q33.1	***HSPD1***	Pure or Complex	Autosomal dominant	(Hansen et al., 2002)
SPG14	3q27-q28		Pure or Complex	Autosomal recessive	(Vazza et al., 2000)
SPG15	14q24.1	***ZFYVE26***	Complex	Autosomal recessive	(Hanein et al., 2008)
SPG16	Xq11.2		Pure or Complex	X-linked	(Steinmuller et al., 1997)
SPG17	11q13	***BSCL2***	Complex	Autosomal dominant	(Windpassinger et al., 2004)
SPG18	8p11.2	***ERLIN2***	Complex	Autosomal recessive or dominant	(Alazami et al., 2011; Rydning et al., 2018)
SPG19	9q33-q34		Pure	Autosomal dominant	(Valente et al., 2002)
SPG20	13q12.3	***SPG20***	Complex	Autosomal recessive	(Patel et al., 2002)
SPG21	15q22.31	***SPG21***	Pure or Complex	Autosomal recessive	(Simpson et al., 2003)
SPG22	Xq13.2	***SLC16A2***	Pure or Complex	X-linked	(Schwartz et al., 2005)
SPG23	1q32.1	***DSTYK***	Complex	Autosomal recessive	(Lee et al., 2017b)
SPG24	13q14		Complex	Autosomal recessive	(Hodgkinson et al., 2002)
SPG25	6q23-q24.1		Complex	Autosomal recessive	(Zortea et al., 2002)
SPG26	12q13.3	***B4GALNT1***	Complex	Autosomal recessive	(Boukhris et al., 2013)
SPG27	10q22.1-q24.1		Complex	Autosomal recessive	(Meijer et al., 2004)
SPG28	14q22.1	***DDHD1***	Pure or Complex	Autosomal recessive	(Tesson et al., 2012)
SPG29	1p31.1-p21.1		Complex	Autosomal dominant	(Orlacchio et al., 2005)
SPG30	2q37.3	***KIF1A***	Pure or Complex	Autosomal recessive or dominant	(Erlich et al., 2011; Lee et al., 2015a)
SPG31	2p11.2	***REEP1***	Pure or Complex	Autosomal dominant	(Züchner et al., 2006)
SPG32	14q12-q21		Complex	Autosomal recessive	(Stevanin et al., 2007b)
SPG33	10q24.2	***ZFYVE27***	Complex	Autosomal dominant	(Mannan et al., 2006)
SPG34	Xq24-q25		Pure	X-linked	(Macedo-Souza et al., 2008)
SPG35	16q23.1	***FA2H***	Pure or Complex	Autosomal recessive	(Edvardson et al., 2008)
SPG36	12q23-24		Complex	Autosomal dominant	(Schule et al., 2009)
SPG37	8p21.1-q13.3		Pure	Autosomal dominant	(Hanein et al., 2007)

Type	Chr	Gene	Phenotype	Inheritance mode	Reference
SPG38	4p16-p15		Complex	Autosomal dominant	(Orlacchio et al., 2008)
SPG39	19p13.2	***PNPLA6***	Complex	Autosomal recessive	(Rainier et al., 2008)
SPG40	3q24-q26		Pure or Complex	Autosomal dominant	(Subramony et al., 2009)
SPG41	11p14.1-p11.2		Pure	Autosomal dominant	(Zhao et al., 2008)
SPG42	3q25.31	***SLC33A1***	Pure	Autosomal dominant	(Lin et al., 2008)
SPG43	19q12	***c19orf12***	Complex	Autosomal recessive	(Landouré et al., 2012)
SPG44	1q42.13	***GJC2***	Complex	Autosomal recessive	(Orthmann-Murphy et al., 2009)
SPG45 and 65	10q24.33	***NT5C2***	Pure or Complex	Autosomal recessive	(Novarino et al., 2014)
SPG46	9p13.3	***GBA2***	Complex	Autosomal recessive	(Martin et al., 2013)
SPG47	1p13.2	***AP4B1***	Complex	Autosomal recessive	(Abou Jamra et al., 2011)
SPG48	7p22.1	***AP5Z1***	Pure or Complex	Autosomal recessive	(Slabicki et al., 2010)
SPG49	14q32.31	***TECPR2***	Complex	Autosomal recessive	(Oz-Levi et al., 2012)
SPG50	7q22.1	***AP4M1***	Complex	Autosomal recessive	(Abou Jamra et al., 2011)
SPG51	15q21.2	***AP4E1***	Complex	Autosomal recessive	(Abou Jamra et al., 2011)
SPG52	14q12	***AP4S1***	Complex	Autosomal recessive	(Abou Jamra et al., 2011)
SPG53	8p22	***VPS37A***	Complex	Autosomal recessive	(Zivony-Elboum et al., 2012)
SPG54	8p11.23	***DDHD2***	Complex	Autosomal recessive	(Schuurs-Hoeijmakers et al., 2012)
SPG55	12q24.31	***C12orf65***	Complex	Autosomal recessive	(Shimazaki et al., 2012)
SPG56	4q25	***CYP2U1***	Pure or Complex	Autosomal recessive	(Tesson et al., 2012)
SPG57	3q12.2	***TFG***	Complex	Autosomal recessive	(Beetz et al., 2013a)
SPG58	17p13.2	***KIF1C***	Pure or Complex	Autosomal recessive	(Novarino et al., 2014)
SPG59	15q21.2	***USP8***	Complex	Autosomal recessive	(Novarino et al., 2014)
SPG60	3p22.2	***WDR48***	Complex	Autosomal recessive	(Novarino et al., 2014)
SPG61	16p12.3	***ARL6IP1***	Complex	Autosomal recessive	(Novarino et al., 2014)
SPG62	10q24.31	***ERLIN1***	Pure	Autosomal recessive	(Novarino et al., 2014)
SPG63	1p13.3	***AMPD2***	Complex	Autosomal recessive	(Novarino et al., 2014)
SPG64	10q24.1	***ENTPD1***	Complex	Autosomal recessive	(Novarino et al., 2014)
SPG66	5q32	***ARSI***	Complex	Autosomal recessive	(Novarino et al., 2014)
SPG67	2q33.1	***PGAP1***	Complex	Autosomal recessive	(Novarino et al., 2014)
SPG68	11q13.1	***FLRT1***	Complex	Autosomal recessive	(Novarino et al., 2014)
SPG69	1q41	***RAB3GAP2***	Complex	Autosomal recessive	(Novarino et al., 2014)
SPG70	12q13.3	***MARS***	Complex	Autosomal recessive	(Novarino et al., 2014)
SPG71	5p13.3	***ZFR***	Complex	Autosomal recessive	(Novarino et al., 2014)
SPG72	5q31.2	***REEP2***	Pure	Autosomal recessive or dominant	(Esteves et al., 2014)
SPG73	19q13.33	***CPT1C***	Pure	Autosomal dominant	(Rinaldi et al., 2015)
SPG74	1q42.13	***IBA57***	Complex	Autosomal recessive	(Lossos et al., 2015)
SPG75	19q13.12	***MAG***	Complex	Autosomal recessive	(Novarino et al., 2014)
SPG76	11q13.1	***CAPN1***	Complex	Autosomal recessive	(Gan-Or et al., 2016)

Type	Chr	Gene	Phenotype	Inheritance mode	Reference
SPG77	6p25.1	***FARS2***	Pure	Autosomal recessive	(Yang et al., 2015)
SPG78	1p36.13	***ATP13A2***	Complex	Autosomal recessive	(Estrada-Cuzcano et al., 2017)
SPG79	4p13	***UCHL1***	Complex	Autosomal recessive	(Rydning et al., 2017)
SPOAN	11q13.2	***KLC2***	Complex	Autosomal recessive	(Melo et al., 2015)
SPG Un	9q22.31	***BICD2***	Pure or Complex	Autosomal recessive or dominant	(Novarino et al., 2014; Oates et al., 2013)
SPG Un	19p13.2	***DNM2***	Pure	Autosomal dominant	(Sambuughin et al., 2015)
SPG Un	9p13.2	***EXOSC3***	Complex	Autosomal recessive	(Halevy et al., 2014)
SPG Un	1p13.3	***KCNA2***	Complex	Autosomal dominant	(Manole et al., 2017)
SPG Un	2p25.1	***KIDINS220***	Complex	Autosomal dominant	(Josifova et al., 2016)
SPG Un	6p21.1	***KLC4***	Complex	Autosomal recessive	(Bayrakli et al., 2015)
SPG Un	1q42.3	***LYST***	Complex	Autosomal recessive	(Shimazaki et al., 2014)
SPG Un	mitochondria	***MT-ATP6***	Complex	Maternally inherited SPG	(Verny et al., 2011)
SPG Un	1p36.22	***MTHFR***	Complex	Autosomal recessive	(Bathgate et al., 2012)
SPG Un	22q13.1	***PLA2G6***	Complex	Autosomal recessive	(Ozes et al., 2017)
SPG Un	19p13.3	***TUBB4A***	Complex	Autosomal dominant	(Kancheva et al., 2015)
SPG Un	9p13.3	***VCP***	Complex	Autosomal dominant	(Bot et al., 2012)
MM_SPG	Xq28	***ABCD1***	Pure	X-linked	(Zhan et al., 2013)
MM_SPG	1q21.3	***ADAR***	Pure	Autosomal Dominant	(Crow et al., 2014)
MM_SPG	18p11.21	***AFG3L2***	Complex	Autosomal recessive or dominant	(Pierson et al., 2011)
MM_SPG	2q33.1	***ALS2***	Pure	Autosomal recessive	(Wakil et al., 2014)
MM_SPG	6q23.2	***ARG1***	Complex	Autosomal recessive	(Baranello et al., 2014)
MM_SPG	5p15.2	***CCT5***	Complex	Autosomal recessive	(Bouhouche et al., 2006)
MM_SPG	22q11.1	***CECR1***	Complex	Autosomal recessive	(Poswar et al., 2016)
MM_SPG	8p23.3	***CLN8***	Complex	Autosomal recessive	(Novarino et al., 2014)
MM_SPG	19p13.2	***DNMT1***	Complex	Autosomal dominant	(Kara et al., 2016)
MM_SPG	3q27.1	***EIF2B5***	Complex	Autosomal recessive	(Novarino et al., 2014)
MM_SPG	6q14.1	***ELOVL4***	Complex	Autosomal recessive	(Aldahmesh et al., 2011)
MM_SPG	5p15.1	***FAM134B***	Complex	Autosomal recessive	(Ilgaz Aydinlar et al., 2014)
MM_SPG	2q31.1	***GAD1***	Complex	Autosomal recessive	(Lynex et al., 2004)
MM_SPG	14q22.2	***GCH1***	Complex	Autosomal dominant or recessive	(Fan et al., 2014)
MM_SPG	6q22.31	***GJA1***	Complex	Autosomal dominant or recessive	(Furuta et al., 2012)
MM_SPG	3p22.3	***GLB1***	Complex	Autosomal recessive	(Kumar et al., 2016)
MM_SPG	4q22.1-q22.2	***GRID2***	Complex	Autosomal dominant	(Maier et al., 2014)
MM_SPG	2q24.2	***IFIH1***	Pure	Autosomal dominant	(Crow et al., 2014)
MM_SPG	9p24.3	***KANK1***	Complex	Imprinting-like inheritance	(Lerer et al., 2005)
MM_SPG	mitochondria	***MT-CO3***	Complex	Maternally inherited SPG	(Tiranti et al., 2000)
MM_SPG	mitochondria	***MT-TE***	Complex	Maternally inherited SPG	(Lax et al., 2013)
MM_SPG	mitochondria	***MT-TI***	Pure or Complex	Maternally inherited SPG	(Corona et al., 2002)
MM_SPG	19q13.32	***OPA3***	Complex	Autosomal recessive or dominant	(Arif et al., 2013)
MM_SPG	13q14.3	***RNASEH2B***	Pure	Autosomal recessive	(Crow et al., 2014)
MM_SPG	13q12.12	***SACS***	Complex	Autosomal recessive	(Gregianin et al., 2013)

Type	Chr	Gene	Phenotype	Inheritance mode	Reference
MM_SPG	20q11.23	***SAMHD1***	Complex	Autosomal recessive	(Dale et al., 2010)
MM_SPG	11p15.4	***TPP1***	Complex	Autosomal recessive	(Kara et al., 2016)

1.2. Clinical diagnosis and heterogeneity

1.2.1 Clinical diagnosis and HSPs differential diagnostics

The HSP diagnosis is established when, in addition to the characteristic clinical symptoms and a neurological examination showing a corticospinal tract involvement, there is a family history of the disease (although sporadic cases are common) or a molecular diagnosis (when the genetic cause is identified). In the absence of family history, it is essential to exclude other potential causes, such as some metabolic or other neurodegenerative disorders since spasticity is a common consequence of multiple disorders. In the complex forms it could be even more difficult to achieve a HSP diagnosis as sometimes the additional neurological symptoms that characterize these forms may appear before the spastic paraplegia itself (Souza et al., 2016). So, additional tests like a cerebral and spinal MRI, electroneuromyography, nerve conduction studies, lumbar puncture with cerebrospinal fluid analysis, long chain fatty acids, white cell enzymes, plasma amino acids, serum lipoprotein analysis, vitamin B12 or vitamin E, copper and ceruloplasmin, serum serology for syphilis, *Treponema pallidum*, human T-cell leukemia virus 1 (HTLV-I) and human immunodeficiency virus (HIV), and a neuro-ophthalmological assessment are usually of help to rule out other causes. The number of differential diagnosis is wide and include, besides other neurodegenerative diseases, also some metabolic, structural, inflammatory, infectious and brain metal accumulation disorders (Table 2) (Parodi et al., 2017; Souza et al., 2016; Tesson et al., 2015).

Table 2: List of some of the main differential diagnosis of HSP.

Differential diagnosis of hereditary spastic paraplegias					
Structural anomalies	**Infectious diseases**	**Inflammatory**	**Metabolic**	**Brain metal accumulation disorder**	**Other neurodegenerative**
- Arnold-Chari malformation - Cerebral palsy - CNS tumor - Spinal cord injury -Cervical Spondylotic myelopathy	- Tropical spastic paraparesis (HTLV-1) -HIV -Neurosyphilis -Neuroborreliosis	- Multiple Sclerosis - Vasculitic Myelopathy - Stiff person syndrome - Sarcoidosis	- Leukodystrophies like: Adrenoleucodystrophy, Pelizaeus-Merzbacher disease, Krabbe disease - Vitamin deficiency (B12, E) - Copper deficiency - Aminoacid disorders like Arginase deficiency - Mitochondrial disorders - Dystonia - Gaucher disease	- Neurodegeneration with brain iron accumulation (NBIA) - Wilson Disease	- Spastic ataxias - Spinocerebellar ataxias - Amyotrophic Lateral Sclerosis (ALS) - Progressive Muscular Atrophy - Progressive Bulbar Palsy - Primary Lateral Sclerosis (PLS) - Inherited dementias

1.2.2 Clinical, genetic and molecular heterogeneity

The large heterogeneity found in HSP impairs both the clinical diagnosis and the genetic identification. Many of the genes already associated to hereditary spastic paraplegia are also associated to several other disorders that seem to clinically overlap. Some examples are *SPG11* associated with amyotrophic lateral sclerosis (ALS) and Charcot-Marie-Tooth (CMT) disease; *KIF1A* (SPG30) associated with hereditary sensory neuropathy (HSN) and mental retardation; *GJC2* (SPG44) associated with hypomyelinating leukodystrophy (HLD); *KIF1C* (SPG58) with spastic ataxia and *ATP13A2* (SPG78) and *UCHL1* (SPG79) were previously linked to Parkinson disease (PD). All of these genes in addition with *SPG7*, *ZFYVE26* (SPG15), *FA2H* (SPG35), *PNPLA6* (SPG39), *GBA2* (SPG46) and *CAPN1* (SPG76) were also reported as causing ataxia (Synofzik and Schüle, 2017). ARSACS is a good example of this clinical overlap between HSP and hereditary ataxia, where variants in the *SACS* gene are responsible for spastic ataxia, but there are reports of patients where leg spasticity or ataxia were absent (Bouhlal et al., 2011; Gregianin et al., 2013). Neurodegeneration with brain iron accumulation (NBIA) is another disease that shares genes with HSP: *C9orf72* and *FA2H*. HSP and the leukodystrophies share at least eleven associated genes (*PLP1*, *HSPD1*, *GJC2, ABCD1, ADAR, EIF2B5, GJA1, IFIH1, RNASEH2B, SAMHD1* and *TUBB4A*). Sometimes, it is the nature of the variant and/or its localization that can impact the phenotype, as illustrated by the inheritance pattern associated with disease-causing variants in *KIF1A*. While loss of function variants are associated with AR neuropathy, missense heterozygous variants in conserved residues of the protein ATPase domain, are responsible for a severe early onset complex AD HSP. *HSPD1* is an example where heterozygous variants were associated with HSP, but homozygous missense variants lead to hypomyelinating leukodystrophy type 4 (Magen et al., 2008). Another example is *TFG* where a heterozygous variant in the P/Q rich domain was implicated in an autosomal-dominant hereditary motor and sensory neuropathy (Ishiura et al., 2012) and a homozygous variant in the coil-coil domain was linked to HSP (Beetz et al., 2013a), which probably indicates the presence of different pathological mechanisms.

HSP-associated genes are so heterogeneous that are constantly being found to be involved in causing a variety of symptoms specific to other disorders, resulting in a huge phenotypic complexity (Figure 1). This can complicate the diagnosis and it will depend on the physician (and the disease evolution at the time of examination) to decide which symptom is most prominent (Tesson et al., 2015). In fact, as the number of reports is growing, the number of genes associated only with pure HSP is decreasing and findings of pure and complex phenotypes have been described for most HSP genes (Klebe et al., 2015; Tesson et al., 2015). Clinical signs of intellectual disability and ataxia are common findings. The identification of peripheral neuropathy is frequent and shows the overlap with Charcot-Marie-Tooth (CMT)

(Klebe et al., 2015). It has already been shown that diseases like spinocerebellar ataxias (SCAs) and CMT seem to share with HSP not only genes and phenotypes but also have overlapping molecular mechanisms (Synofzik and Schüle, 2017; Timmerman et al., 2013). Currently, the picture presents more like a phenotypic spectrum than rigid different clinical entities. Thus, when genetically studying (mainly the complex) HSP families, we should be aware that our answer could be in a gene associated with one of these related disorders: ALS, CMT, SCAs, PLS, leukodystrophies and NBIA.

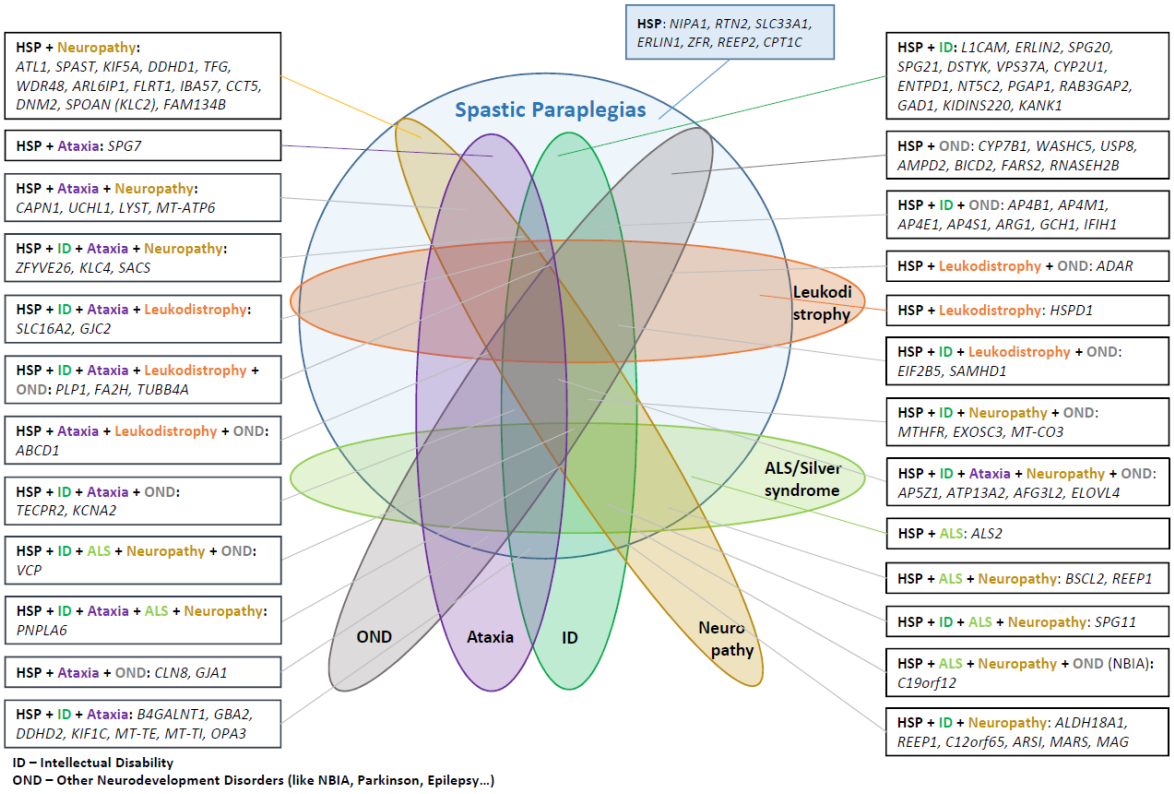


Figure 1: Overview of the phenotypic complexity associated to HSP genes

1.3. Mechanisms of pathogenesis in the HSPs

The molecular pathways leading to the disease, as the clinics and genetics of HSPs, also appear to be heterogeneous. The common hallmark of the disease is the axonal degeneration maximal in the longest tracts. Postmortem analysis of patient tissues consistently identified axonal degeneration with maximum severity at the distal ends of the corticospinal tracts and fasciculus gracilis fibers (long descending and ascending pathways of the spinal cord) and of the spinocerebellar tracts. These tracts constitute the longest motor and sensory axons of the central nervous system (CNS). The degeneration seems to be mostly limited to the cell axon,

starting in the distal ends of the axons and progressing toward the cell body, in what led to the classification of HSP as a primary CNS “dying back” axonopathy (Crosby and Proukakis, 2002; DeLuca et al., 2004; Schwarz and Liu, 1956). The pathophysiological mechanism(s) identified should so elucidate why the longest neurons in the spinal cord are affected in this disease.

The identification of more than 60 HSP causing genes has been essential for the identification of several pathophysiological mechanisms that could underlines the disease and that sometimes are interrelated. Although many of these genes encode proteins that remain without an identified function and with much to uncover regarding their physiological role, still, it was possible to identify the main common pathways affected by variants in these genes. It has been shown that several of these proteins interact being part of a large interconnected network, which explains why variants in some genes are capable of affecting different pathways such as *SPAST* (Figure 2).

One of the first mechanisms to be related to HSP was the disruption of the axonal transport, which was supported by the identification of several genes associated with the disease that encode kinesin chains (*KIF5A* and later, *KIF1A, KIF1C and KLC2*), an essential part of the anterograde microtubule-dependent axonal transport (Crosby and Proukakis, 2002; Reid et al., 2002). Other identified mechanisms are the disruption of membrane trafficking, mitochondrial function, ER morphogenesis, lipid metabolism, myelination, autophagy, DNA repair and nucleotide metabolism.

These mechanisms can explain why the disease affects primarily these long axons because, as a result of their length (that can reach up to 1 meter), and unique and highly polarized architecture, they are really dependent and sensitive to distress of numerous cellular processes as axonal transport, membrane trafficking, cytoskeletal dynamics, endoplasmic reticulum morphogenesis and endocytic recycling. They require that these complex machineries work properly to be able to efficiently drive the transport of signals, molecules, and organelles, so they are also dependent of a capable mitochondrial function (Blackstone et al., 2011; Crosby and Proukakis, 2002; Salinas et al., 2008). Several of these pathways were also associated with other neurodegenerative disorders like SCAs and ALS, showing once more the interconnection between these disorders.

1.3.1 Axonal transport and microtubule dynamics

As HSP mainly affects the longest neurons in the spinal cord, which are extremely dependent on a correct transport and delivery of cell components, it was early hypothesized that defects in the axonal transport could be the main pathophysiological mechanism leading to the disease.

Axonal transport depends mainly on polarized microtubules. The anterograde motility (moving towards the axonal terminals) is powered by kinesin motor proteins and is responsible for the transport of organelles like mitochondria, lysosomes, and endosomes as well as vesicles, RNA and newly synthesized proteins and lipids. Also, synaptic components necessary for the pre-synaptic activity, short microtubules and neurofilaments required for the maintenance of the microtubule tracks are transported this way. The retrograde motility (moving toward the cell body) is driven by dynein motor protein and is essential to maintain the neuron homeostasis because it clears the distal ends of the axons of aging proteins and organelles directing them for degradation. It has also a role in delivering neurotrophic and injury-generated signals back to the cell bodies (Baas et al., 2006; Maday et al., 2014).

Up to now, at least 8 HSP-associated proteins were linked to axonal transport and microtubule dynamics (Figure 2 and Supplementary Table 1). *SPAST* (SPG4) was one of the first identified HSP causing genes and is the most common cause of the disease. It encodes spastin, a protein shown to be involved in microtubule dynamics as it appears to be a microtubule-severing protein. Spastin has three main structural domains required for the microtubule breakage: a microtubule interacting and endosomal trafficking (MIT) domain, a second microtubule binding domain (MBD) responsible for the binding of spastin to the microtubules and at the C-terminus, a catalytic AAA ATPase domain that contains the enzymatic activity necessary for microtubule severing (Blackstone et al., 2011; Errico et al., 2002; Salinas et al., 2007). Most of the *SPAST* disease-causing variants found are localized in this AAA domain and seem to act by a haploinsufficiency effect (Shoukier et al., 2009). *SPG20* seem to share with *SPG4* this conserved MIT domain, suggesting also an effect of this protein in microtubule binding and in intracellular transport (Ciccarelli et al., 2003). *TUBB4A* was also linked to microtubule dynamics as it encodes β-tubulin, one of the major subunits of the microtubules. Deleterious variants in this gene lead to microtubule accumulation in a rat model (Duncan et al., 2017a).

Later, the identification of causative variants in *KIF5A* that codes for a kinesin heavy chain, was a major evidence for the role of an abnormal axonal transport as a key pathophysiological mechanism in HSPs. Kinesin heavy chain is an integral part of the motor protein involved in anterograde axonal transport. Variants in *KIF5A* are mostly localized in the extremely conserved motor domain. Functional studies showed that these variants in *KIF5A* lead to a reduced microtubule affinity and/or gliding velocity, which leads to a deficiency in the axonal transport of kinesin-dependent cargoes and ultimately to axonal degeneration (Ebbing et al., 2008; Salinas et al., 2008). Disease-causing variants in another HSP causing genes coding for kinesins chains were later discovered (*KIF1A, KIF1C* and *KLC2*) increasing the evidence for the role of axonal transport defects for the disease. *BICD2*, a gene responsible for spinal muscular atrophy, which was associated with HSP, is also linked to problems in the axonal

transport as it is involved in dynein-mediated transport. *BICD2* seems to be responsible for linking the dynein motor complex to several cargoes including Golgi vesicles, lipid droplets, nuclei, and specific mRNAs (Neveling et al., 2013; Teuling et al., 2008).

1.3.2 Endoplasmic reticulum morphogenesis

The endoplasmic reticulum (ER) is a multifunctional organelle involved in several essential tasks for the cell, being the most abundant membrane compartment. It is responsible for the synthesis, modification, quality control and trafficking of secreted and integral membrane proteins. It has a role in the Ca^{2+} regulation being the major Ca^{2+} storage compartment of the cell. It is critical for the lipid and sterol biosynthesis and distribution, carbohydrate metabolism, peroxisome biogenesis and for the formation of enzymes involved in drug detoxification (Blackstone, 2012; Lynes and Simmen, 2011; Schwarz and Blower, 2016). Additionally, ER interacts with several other membranes, including mitochondria, the Golgi apparatus, the plasma membrane and with lipid droplets (LDs) (Raiborg et al., 2015). The ER, due to its multitude of roles, it is also very heterogeneous in its form, with the distinct morphologies seeming to correlate with its functional specializations. It is a continuous membrane system that includes the nuclear envelope, sheet-like cisternae, and a network of interconnected tubules extending into the cell periphery. Formation and maintenance of ER morphology depend on the proper segregation of diverse types of shaping proteins to generate the highly-curved ER tubules and the fusion of tubules to produce the three-way junctions required to obtain a network. Additionally, it is the interaction among cytoskeletal elements, plasma membrane, and organelles as well as numerous signaling pathways that cooperate to dynamically position and shape the ER (Goyal and Blackstone, 2013; Schwarz and Blower, 2016).

The importance of a proper ER morphology in highly-polarized cells such as neurons was recognized with the discovery of several HSP-causing genes involved in shaping the ER network which included *ATL1*, *SPAST*, *RTN2* and both *REEP1* and *REEP2* genes. Causative variants in *ATL1* (SPG3A) are the second most frequent cause of AD HSP. This gene encodes the protein atlastin-1, a member of the superfamily of dynamin-related GTPases that localize predominantly to the tubular ER and mediate the fusion of ER tubules (Hu et al., 2009; Orso et al., 2009). *RTN2* (SPG12) encodes reticulon 2, which also localizes to the ER tubules and that acts together with DP1/Yop1 as curvature-stabilizing proteins (Voeltz et al., 2006). *REEP1* (SPG31) and *REEP2* (SPG72) are part of the REEP/DP1/Yop1p family of ER-shaping proteins, which as RTN2 protein has two hydrophobic domains that, when inserted into the phospholipid bilayer, are thought to form hairpin loops generating a curvature in the ER

membrane. Spastin interacts with atlastin-1, REEP1 and reticulon 2 proteins and they all seem to contribute to the shaping of the tubular ER network (Beetz et al., 2013b; Esteves et al., 2014; Shibata et al., 2010; Voeltz et al., 2006). Additionally, REEP1, spastin and atlastin-1 interact within the tubular ER, seeming to be responsible for mediating the interaction of the tubular ER with microtubules (Park et al., 2010). The microtubule-based motility in the ER is important for the proper organization and distribution of the ER tubules and the impairment of this link between the tubular ER and the microtubule cytoskeleton seem to be one of the HSP pathogenic mechanisms (Blackstone, 2012). Other genes which proteins are associated with ER morphogenesis are *TFG* (SPG57), *ARL6IP1* (SPG61) and *RAB3GAP2* (SPG69). TFG protein is important for the organization of transitional ER and ER exit sites into larger structures and variants in this gene lead to a disruption of the peripheral ER tubules organization and collapse of the ER network (Beetz et al., 2013a; Elsayed et al., 2017; McCaughey et al., 2016). ARL-6-interacting protein 1 has a reticulon-like short hairpin transmembrane domains and seems to, as reticulon 2, be able to shape high-curvature ER tubules and to bind atlastin-1 (Yamamoto et al., 2014). *RAB3GAP2* encodes a protein of the Rab regulatory complex, which localizes to the ER and is responsible for activating RAB18 that mediates its recruitment to the ER. Their absence results in altered ER structure and it seems that RAB18 effectors may regulate ER tubule tethering and fusion or maybe work as inhibitors of ER sheet extension (Gerondopoulos et al., 2014). *BSCL2* (SPG17), *ERLIN1* (SPG62) and *ERLIN2* (SPG18) encode proteins that also localize to the ER and relate to both ER and lipid metabolism pathways due to the key function of the ER in the synthesis, metabolism and distribution of lipids and sterols (Blackstone, 2012). *BSCL2* encodes an ER protein thought to act at the interface of the ER with LDs regulating their size. Missense causative variants seem to result in misfolding of the protein leading to aggregate formation, increased level of ER stress-mediated molecules and induced apoptosis in cultured cells (Ito and Suzuki, 2007; Salo et al., 2016). *ERLIN2* (SPG18) and *ERLIN1* (SPG62) encode proteins that localize preferentially to cholesterol-rich domains including lipid rafts and are involved in the regulation of the cellular cholesterol homeostasis. Additionally, ERLIN proteins are associated to ER-associated degradation (ERAD) that regulates protein degradation by the ubiquitin-proteasome machinery (Browman et al., 2006; Huber et al., 2013).

1.3.3 Membrane trafficking

Cells depend on a highly regulated and dynamic membrane trafficking system for a proper intracellular distribution of proteins, lipids and complex carbohydrates. This transport relies on membrane bound vesicles and can occur within different organelles inside the cell, or through

the plasma membrane to and from the extracellular environment (Herrmann and Spang, 2015). Traffic of cargoes occurs in two different directions, to the plasma membrane, endosomes, and lysosomes or out of the cell (exocytosis), which is the direction of the secretory pathway or, in the opposite direction (endocytosis). In the secretory pathway, newly synthesized proteins, lipids and carbohydrates are transported from the ER via the Golgi to their final destination. In the opposite direction and as an important part of endocytosis is the degradative pathway via autophagy to clear the cell of damaged and unwanted proteins. Exocytosis and endocytosis use vesicles that enclose the cargoes and transport them from one membrane to another. Briefly, the steps that occur are: 1) cargo selection and vesicle budding from a donor compartment, which is mediated by protein coats, 2) vesicle transport along the microtubule cytoskeleton, 3) vesicle targeting and fusion with an acceptor compartment, which is dependent of a machinery that includes the Rab family of small GTPases and the SNARE proteins. There are at least 3 types of well characterized coat proteins: i) clathrin-coated vesicles that mediate endocytosis toward the endosomes and the Golgi; ii) coat protein I (COPI) that surround vesicles to mediate transport within the Golgi and towards the ER; iii) COPII-coated vesicles that mediate transport in the opposite direction, the secretory pathway from the ER towards the Golgi. For the last step, to allow the cargo to be delivered to the exact location, members of the Rab small GTPase family are needed to provide specificity, mediating the intracellular destination of the vesicles. For the dock of the vesicles to the correct location and fusion with membranes, SNARE proteins are required (Bonifacino and Glick, 2004; Herrmann and Spang, 2015; Salinas et al., 2008).

Endosomes are a big part of this traffic as they are organelles that function as a central sorting site for the secretory and degradative pathways, and with which the endocytic vesicles fuse to release their content. The endosomal system is also important for retargeting membrane components as it is able to internalize and degrade mislocalized proteins or recycle them from one domain to another. There are three main pathways that interact with endosomes: (1) vesicles constantly moving between the Golgi and endosomes in both directions, (2) between the plasma membrane and early endosomes (through recycling endosomes) in both directions and (3) late endosomes fusing with lysosomes. Endosomes mature into late endosomes and when fused with a lysosome, the internalized vesicles are delivered to the lysosome lumen for degradation. Proteins are marked for this pathway by ubiquitin, which is used by the "endosomal sorting complexes required for transport" (ESCRT) to sort the targeted protein for ubiquitin-dependent degradation (Crosby and Proukakis, 2002; Lasiecka and Winckler, 2011; Noreau et al., 2014).

Numerous HSP proteins are implicated in membrane trafficking, from regulators of the vesicle and endosomal trafficking, to proteins involved in the morphogenesis of the several organelles involved. *ATL1* (SPG3A) encodes a dynamin-related GTPase, which is present both in ER and

Golgi compartments. Disease-causing variants in this gene lead not only to altered vesicle trafficking between ER and Golgi but also to altered Golgi morphogenesis (Namekawa et al., 2007). *AP4B1* (SPG47), *AP4M1* (SPG50), *AP4E1* (SPG51), and *AP4S1* (SPG52) encode proteins that constitute the adaptor protein 4 (AP-4) complex, which mediates vesicle formation and sorting of cargo for vesicle inclusion in the secretory and endocytic pathways (Abou Jamra et al., 2011; Hirst et al., 2013a). *AP5Z1* (SPG48) encodes one of the subunits of the adaptor protein 5 (AP-5) complex, which was also found to be involved in endosomal sorting and interacts with spatacsin (*SPG11*) and zinc finger FYVE domain-containing protein 26 (SPG15) that seem to be present in ER, protein-trafficking vesicles and microtubules and are involved in the recycling of lysosomes (Chang et al., 2014; Hirst et al., 2011; Murmu et al., 2011; Renvoisé et al., 2014). Spastin (SPG4) in addition to be involved in microtubule dynamics and ER morphogenesis, its MIT domain (microtubule interacting and endosomal trafficking domain) binds to the ESCRT system responsible for recognizing ubiquitin-modified proteins. Spartin (*SPG20*) also contains an MIT domain and selectively interacts with a component of the ESCRT-III complex (Blackstone, 2012). Magnesium transporter NIPA1 (SPG6) is a neuron-specific transmembrane protein expressed in early endosomes, cell surface and in the Golgi. It co-localizes with atlastin-1, and were characterized as binding partners, with knock-down of *NIPA1* expression resulting in a marked reduction of atlastin-1 in neuronal processes (Botzolakis et al., 2012). A rat model with a *NIPA1* causative variant showed accumulation of tubular-vesicular organelles with endosomal features starting at axonal and dendritic terminals, followed by multifocal vacuolar degeneration in the CNS and peripheral nerves (Watanabe et al., 2013). Additionally, *NIPA1, ATL1, SPAST* and *SPG20* are involved in the regulation of bone morphogenetic proteins (BMP) signaling, crucial for axonal growth and synaptic function among other developmental processes (Blackstone et al., 2011; Lo Giudice et al., 2014). *WASHC5*, formerly known as *KIAA0196* (SPG8) encodes a subunit of the large protein complex WASH, which connects endosomes to the cytoskeleton, through an interaction with vacuolar protein sorting-associated protein 35 (VPS35) (Harbour et al., 2010). Strumpellin (SPG8) also interact with valosin-containing protein (VCP), an AAA ATPase also recently related with HSP (Clemen et al., 2010). Vacuolar protein sorting-associated protein 37A (SPG53) encodes a component and a upstream regulator of ESCRT-I, another member of the endosomal sorting complex required for the ESCRT system (Bache et al., 2004). Maspardin (*SPG21*) co-localizes with vesicles of the intracellular endosomal/trans-golgi system (Zeitlmann et al., 2001). Phospholipase DDHD1 (SPG28) and phospholipase DDHD2 (SPG54) are two enzymes making the link between membrane sorting and lipid metabolism: DDHD2 localizes to the cis-Golgi and also to the ER-Golgi intermediate compartment suggesting that it facilitates membrane and vesicle fusion by the modification of membranes through phospholipid hydrolysis (Inoue et al., 2012; Schuurs-Hoeijmakers et al., 2012; Yadav and

Rajasekharan, 2016). TFG (SPG57) is required for COPII recruitment and protein secretion from the ER (Witte et al., 2012). Ubiquitin carboxyl-terminal hydrolase 8 (SPG59) has diverse roles in membrane trafficking, from endosomal regulation to retrograde transport (Smith et al., 2016). WD repeat-containing protein 48 (SPG60) is a regulator of deubiquitinating complexes and it seems to be involved with vesicular transport or membrane fusion events necessary for transport to lysosomes in case of infection by Herpesvirus saimiri tyrosine kinase-interacting protein (Park et al., 2002). Cation-transporting ATPase 13A2 recently identified as SPG78 is a lysosomal P5-type transport ATPase and variants in this gene seem to lead to a combination of lysosomal and mitochondrial dysfunction (Estrada-Cuzcano et al., 2017). *DNM2* encodes dynamin 2, a highly conserved large GTPase involved in the release of vesicles during endocytosis and intracellular membrane trafficking (Bitoun et al., 2009). Lysosomal-trafficking regulator (LYST) seems to play a role in homotypic fusion of different prelysosomal organelles and autophagosomes with variants in this gene leading to the accumulation of these organelles (Rahman et al. 2012). Protein bicaudal D homolog 2 (BICD2) encodes a cargo adaptor protein which interacts with Rab6 primarily involved in retrograde transport between the Golgi and ER (Heffernan and Simpson, 2014; Matsuto et al., 2015).

1.3.4 Mitochondrial function

Mitochondria is a very important organelle that in addition to its role in producing ATP, is involved in the response mechanisms to oxidative stress, apoptosis and are a source of toxic free radicals. Mitochondria are present at a high number in neurons, which explain their particular vulnerability to mitochondrial defects with mitochondrial impairment being found in several neurodegenerative diseases like ALS, Alzheimer´s and Parkinson´s diseases (Filosto et al., 2011).

In HSP, at least 10 genes were already associated with mitochondria dysfunction. *SPG7* encodes a protein that together with *AFG3L2* (another HSP-associated gene) forms the m-AAA metalloprotease complex of the inner mitochondrial membrane. This protease functions in protein quality control and is necessary for the regulation of ribosomal assembly (Atorino et al., 2003; Nolden et al., 2005). Fibroblasts from SPG7 patients showed reduced activity of complex I in the mitochondrial respiratory chain and increased sensitivity to oxidative stress with increased mitochondrial DNA damage (Wedding et al., 2014). In a mouse model for SPG7, mitochondrial morphological abnormalities were observed in the synaptic terminals and distal regions of axons correlating with the onset of motor impairment and prior to axonal swelling and degeneration. The axonal swelling occurred due to a huge accumulation of organelles and neurofilaments, suggesting impairment of anterograde axonal transport, which may indicate

that the axonal transport impairment could be secondary to mitochondrial dysfunction (Ferreirinha et al., 2004; Rugarli and Langer, 2006). *HSPD1* (SPG13) encodes a mitochondrial protein known as heat shock protein 60, thought to assist in the folding of a subset of proteins in mitochondria. Causative variants in this gene also lead to impaired mitochondrial quality control by a reduced expression of the mitochondrial quality control proteases Lon (ATP-dependent protease La) and ClpP (ATP-dependent Clp protease proteolytic subunit) (Bross et al., 2008; Hansen et al., 2008). Spartin (SPG20) seems to co-localize with mitochondria with disease-causing variants leading to a loss of this association. Due to the suspected role of this protein in microtubule binding, it was suggested a possible function for this protein in the microtubule-mediated transport of mitochondria (Lu et al., 2006). Receptor expression-enhancing protein 1 (SPG31) also seems to co-localize with mitochondria and a reduction in mitochondrial respiration and complex IV activity were found both in patients muscles and fibroblasts (Goizet et al., 2011; Züchner et al., 2006). *C12orf65* (SPG55) codes for a mitochondrial matrix protein which, when mutated, impairs the mitochondrial protein synthesis and the respiratory complex enzyme activity (Shimazaki et al., 2012). *IBA57* (SPG74) encodes a mitochondrial protein required for [4Fe-4S] protein maturation (Lossos et al., 2015; Sheftel et al., 2012). *ATP13A2* (SPG78) encodes a lysosomal P-type ATPase, which regulates mitochondrial bioenergetics through macroautophagy. Disease causing variants in this gene led to decreased ATP synthesis rates, increased mitochondrial DNA levels, a higher frequency of mitochondrial DNA lesions, increased oxygen consumption rates, and increased fragmentation of the mitochondrial network (Grünewald et al., 2012; Gusdon et al., 2012). *MT-ATP6* encodes a subunit of ATP synthase (Complex V), the final enzyme of the oxidative phosphorylation (OXPHOS) system. It is encoded by the mitochondrial DNA (mtDNA), being thus maternally inherited (Verny et al., 2011). *DDHD1* (SPG28) and *CYP2U1* (SPG56) encode two enzymes involved in lipid metabolism, which seem to have a deleterious impact on mitochondrial bioenergetic function with patients fibroblasts showing alteration of mitochondrial architecture and bioenergetics with increased oxidative stress (Tesson et al., 2012).

1.3.5 Lipid metabolism

Lipid metabolism was recently identified as a mechanism highly affected in HSP. Lipids play many different roles from storing energy, to electrical insulation in myelin, modulation of cellular calcium signaling and other signal transduction cascades. Additionally, are important structural components of the cell membranes. The composition of the membranes in lipids and sterols controls membrane fluidity, as well as compartmentalization into specialized microdomains and is important for crucial processes like exocytosis and ion channel functions. Considering

the high number of intracellular organelle membranes present in neurons and the importance of these membrane properties, it is not surprising that defects in lipid metabolism have been linked to several neurodegenerative disorders, including HSPs. Sterols (such as cholesterol), glycerophospholipids and sphingolipids are the main components of the lipid portion in the brain with its distribution varying between axons (white matter), nerve cell bodies (gray matter) and myelin. Free cholesterol is a crucial structural component of the plasma membrane of a cell and, in humans, the average level of cholesterol in the CNS is higher than in other tissues. Regulation of the cholesterol homeostasis is very important because there is little or no transfer of cholesterol across the blood-brain barrier, so most cholesterol is produced locally in the brain, and there is a precise balance between the biosynthesis, storage, and catabolism of cholesterol metabolites (Klebe et al., 2015; Lim et al., 2012; Tsaousidou et al., 2008).

The first gene to be associated with lipid metabolism was *CYP7B1* (SPG5), which codes for a 7α-hydroxylase involved in cholesterol metabolism. More precisely, it works in two pathways: the ‘‘alternate/ acidic pathway’’ and in the conversion of the cholesterol derivate dehydroepiandrosterone (DHEA) into 7-OH-DHEA. Additionally, there is growing evidence to suggest that this DHEA-related neurosteroids act as neuroprotective factors and can reduce ischemia-induced neurodegeneration, which could relate *CYP7B1* disease-causing variants with a loss of this neuroprotective role (Tsaousidou et al., 2008). In addition, the enzymatic blockage of the pathway by the mutated CYP7B1 enzyme led to the accumulation of 25- and 27-hydroxycholesterol in cerebrospinal fluid and serum, which were demonstrated to be toxic to neurons in cultures (Schöls et al., 2017).

Spartin (SPG20) as seipin (SPG17) (mentioned in 1.3.2) also is involved in lipid droplets (LDs) turnover (Eastman et al., 2009). *B4GALNT1* (SPG26) encodes an enzyme involved in the biosynthesis of complex gangliosides that could lead to alteration of the distribution of phospholipids and cholesterol in membranes (Ohmi et al., 2011). *DDHD1* (SPG28) and *DDHD2* (SPG54) code for enzymes of the phospholipase A1 family which hydrolyze phospholipids producing 2-acyl-lysophospholipids and fatty acids (Gonzalez et al., 2013). *FA2H* (SPG35) encodes the enzyme fatty acid 2-hydroxylase, highly expressed in brain, which catalyzes the 2-hydroxylation of myelin galactolipids, galactosylceramide, and sulfatide (Edvardson et al., 2008). *PNPLA6* (SPG39), known also as *NTE* (neuropathy target esterase) encodes a phospholipase, which play a critical part in lipid membrane homeostasis and influences membrane fluidity (Sogorb et al., 2016). *SLC33A1* (SPG42) encodes an acetyl-CoA transporter, which carry acetyl-CoA into the Golgi apparatus. Acetyl-CoA is essential for maintaining the balance between carbohydrate and fat metabolism and an inadequate supply of acetyl-CoA, can result in misprocessing of gangliosides and glycoproteins (Blackstone, 2012; Lin et al., 2008). Non-lysosomal glucosylceramidase (SPG46) is an enzyme of the sphingolipid metabolism acting on glucosylceramides (Martin et al., 2013). *CYP2U1* (SPG56)

codes for a P450 hydroxylase which catalyze the hydroxylation of arachidonic acid and related long-chain fatty acids (Tesson et al., 2012). Novarino identified two new HSP genes linked to lipid metabolism: *ARSI* (SPG66) which codes a sulfatase that hydrolyzes sulfate esters and sulfamates, and *PGAP1* (SPG67) that encodes an enzyme involved in the biosynthesis of glycosylphosphatidylinositols, which are glycolipid anchors responsible for anchoring several proteins to the cell surface (Murakami et al., 2014; Novarino et al., 2014). *CPT1C* (SPG73) encodes a member of the carnitine/choline acetyltransferase family and it was found that causative variants in this gene impairs the formation of LDs (Rinaldi et al., 2015).

1.3.6 Myelination and development

Myelin is a lipid-rich organelle that surrounds axons, which is constituted of compact membrane layers made and organized by the surrounding oligodendroglia cells. It works as an electrical insulator for axonal membranes and constitutes the white matter tracts in the nervous system. Myelin is composed by a high proportion of lipids (70%–85%) and only 15%–30% of proteins, which presumably contributes to its high electrical resistance, as most cell membranes only contains 30–50% lipids (Alderson et al., 2006; Edvardson et al., 2008). It is so, expected that many of the genes related to the lipid metabolism pathway are also related to myelination as is the case of *B4GALNT1, FA2H* and *PGAP1*. *PLP1* (SPG2) encodes the proteolipid protein, which together with its smaller isoform DM20 are the most abundant myelin proteins in the CNS. The two proteins are involved in myelin sheath stabilization, axon–myelin interactions, and also in oligodendrocyte maturation (Inoue, 2005). *SLC16A2* (SPG22) encodes the monocarboxylate transporter 8 (MCT8), a transporter of the thyroid hormone, which is essential for the CNS development and specifically for oligodendrocytes (La Piana et al., 2015; Younes-Rapozo et al., 2006). *GJC2* (SPG44) codes the gap junction protein connexin 47, which form channels that appear to be crucial for the proper maintenance of myelin (Orthmann-Murphy et al., 2009). *MAG* (SPG75) encodes a “myelin-associated” glycoprotein that functions in glia-axon interactions, involved both in the maintenance of myelinated axons and in the differentiation, maintenance and survival of oligodendrocytes (Quarles, 2007). *L1CAM* (SPG1), one of the first HSP genes identified is not associated with myelin formation/maintenance but with the development and maintenance of the nervous system. It encodes the L1 cell adhesion molecule, a transmembrane glycoprotein of the immunoglobulin super-family expressed mainly in neurons and Schwann cells, that seems to play an important role in the development of the nervous system by promoting neuron-neuron adhesion and axon outgrowth and is involved in guidance of neurons in the developing CNS (Hortsch, 2000; Jouet et al., 1994; Kenwrick et al., 2000). Patatin-like phospholipase domain-

containing protein 6 (SPG39), in addition to the role in lipid metabolism, also has a role in axonal and synaptic integrity being necessary for a normal nervous system development (Chang and Wu, 2010; Glynn, 2013).

1.3.7 Nucleotide metabolism and DNA repair

Recently, Novarino implicated, for the first time, defects in nucleotide metabolism as a pathway to HSP with the identification of variants in *ENTPD1, NT5C2* and *AMPD2* in HSP patients (Novarino et al., 2014). Maintenance of nucleotide equilibrium is critical for all cells, but is especially important in the nervous system, where there is a great necessity for ATP. Nucleotide imbalance can lead to mitochondrial dysfunction as mitochondria are particularly prone to oxidative stress-associated DNA damage. DNA damage triggers an increase in deoxynucleotide triphosphate (dNTP) levels, necessary for DNA repair mechanisms that require unscheduled DNA synthesis (Fasullo and Endres, 2015).

The three newly identified genes *AMPD2* (SPG63), *ENTPD1* (SPG64) and *NT5C2* (SPG45/65) are involved in purine nucleotide metabolism. *AMPD2* encodes the adenosine monophosphate deaminase 2, an enzyme involved in the maintenance of cellular guanine nucleotide pools (Akizu et al., 2013). *ENTPD1* encodes an ectonucleoside triphosphate diphosphohydrolase-1 responsible for the ATP hydrolysis in order to generate extracellular adenosine (Nardi-Schreiber et al., 2017). *NT5C2* encodes a cytosolic 5' nucleotidase, which catalyzes the hydrolysis of adenosine monophosphate and inosine monophosphate releasing adenosine, and is essential for the equilibrium maintenance of the purine's pools in the brain and spinal cord (Elsaid et al., 2017).

WD repeat-containing protein 48 (SPG60) beyond being involved in vesicular transport it was also associated with DNA repair mechanisms, interacts with ubiquitin carboxyl-terminal hydrolase 1, a protein with deubiquitylating activity, activating the catalytic activity of the USP1 complex. This addition and removal of ubiquitin or ubiquitin-like proteins from factors is a key mechanism for the modulation of the DNA damage response pathways (Jacq et al., 2013).

1.3.8 Autophagy

Autophagy is a complex process responsible for the transport and degradation of cytoplasmic components in the lysosome, which include compromised proteins and organelles. It is a major intracellular mechanism necessary to maintain cellular homeostasis in all cells, however

neurons are amongst the most sensitives to autophagy dysfunction (Kochergin and Zakharova, 2016; Oz-Levi et al., 2013).

TECPR2 (SPG49) brought a new pathway to the HSPs, linking autophagy with the disease. It encodes a member of the tectonin β-propeller repeat-containing family, containing both TECPR and tryptophan-aspartic acid repeat (WD repeat) domains. This gene has been implicated in autophagy due to its similarity with tectonin beta-propeller repeat-containing protein 1 (TECPR1) and Hermansky Pudlak syndrome 5 protein (HPS5) (both linked with the autophagic pathway), and as reduced expression levels of this gene seem to lead to impaired autophagy (Oz-Levi et al., 2012). *SPG11* and *ZFYVE26* (SPG15) encoded proteins are also linked to autophagy due to their role in autophagic lysosome reformation (Chang et al., 2014). Acetyl-coenzyme A transporter 1 (SPG42) seems to be involved in the regulation of autophagy induction (Pehar et al., 2012) and finally, cation-transporting ATPase 13A2 (SPG78) is also involved in the regulation of the autophagy-lysosome pathway by affecting synaptotagmin-11 (PD-associated gene) levels (Bento et al., 2016).

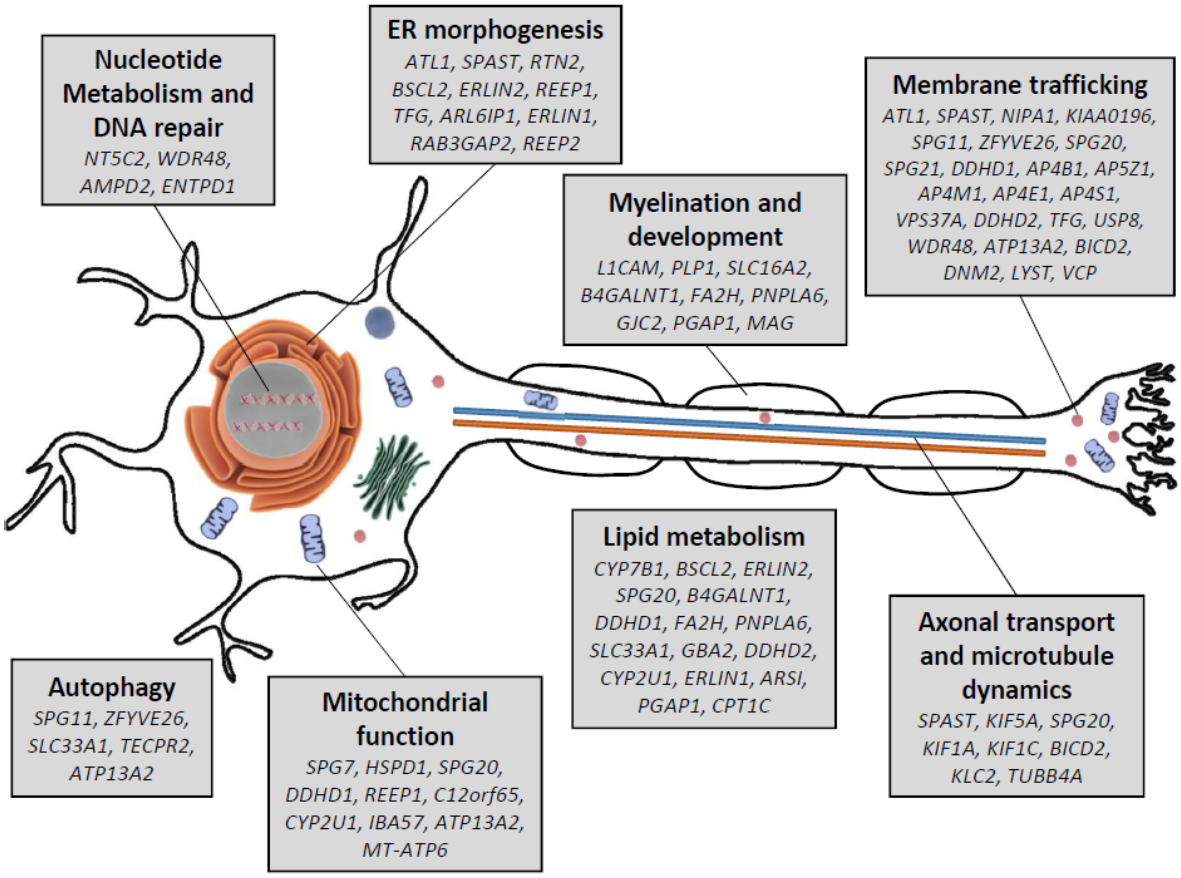


Figure 2: Illustrative representation of the pathophysiological mechanisms and related genes involved in HSP

1.4. Animal models

1.4.1 HSP animal models

Hereditary spastic paraplegias with its large diversity of clinical, genetic and molecular pathways led to the creation of innumerous animal models for the study of the disease, since 1997 until now (Branchu et al., 2017; Dahme et al., 1997). Animal models of HSP include, ordered by frequency: mouse, zebrafish, *Drosophila*, *C. elegans*, and rat. There are also natural models of the disease in mouse, dog and cattle. The creation of animal models for these many HSP genetic types is expected to allow the thorough exploration of the involved disease mechanisms for each gene and the potential trial of candidate treatments in a more reliable way as the profound complexity of the neural networks cannot be easily reproduced "in vitro". Also, with the constant identification of new genes and variants by NGS, the possibility of creating transient models allows a molecular confirmation of the findings pathogenicity (Fink, 2013; LeDoux, 2005; Patten et al., 2014).

HSP animal models should be able to produce some disease features, such as age-dependent motor impairment and axon degeneration and also allow the evaluation of genetic mechanisms like haploinsufficiency, dominant-negative or toxic gain-of-function. Although no animal model can reproduce all of the subtleties of a human disease as there are evidences suggesting some species-specific effects (Thomsen et al., 2010), still these genetic models do reflect important cellular, molecular and physiological features pertinent to disease biology and pharmacological screening (Fink, 2013; Rainier and Fink, 2005). The choice of the appropriate animal model should take into consideration the aim of the experiment and the species characteristics. All have different life spans, developmental periods until maturity and specific advantages. Even within species, strain selection can bring different compensations and may be critical for the success of an experiment. *Caenorhabditis elegans* is a great organism for studies such as large-scale RNA interference (RNAi) experiments as they fed on bacteria that could be expressing any target-gene dsRNA, are inexpensive, have a short development period, are translucent, can be maintained in 96-well plates for high-throughput analyses and genetic manipulations and also can be frozen for long-term storage. *Drosophila melanogaster* is also inexpensive and with a short developmental period and as it has been used for a long time in genetics, there are several powerful genetic tools and numerous online resources available for modeling human disorders in *Drosophila*. *Danio rerio* (Zebrafish) is an inexpensive vertebrate very useful for developmental genetics as it has the vertebrate physiological systems (e.g., neural, cardiovascular, gastrointestinal), and the embryos are translucent. Mice (*Mus musculus*) are probably the most commonly used mammalian model due to their genetic similarity to humans, ease of maintenance and handling, and their high

reproduction rate. It is mainly useful in the context of neurodegenerative and movement disorders research where age-associated changes and motor function could be more easily studied. Rat models (Rattus norvegicus) are useful for physiological studies as they are large enough for the study of most diseases relevant to humans (LeDoux, 2005).

1.4.2 Zebrafish model

Zebrafish has been increasingly becoming a valued and respected model being now used as a model for a wide variety of human diseases. This species, as a vertebrate, has a similar overall nervous system organization to humans and as also possess human homologue genes with conserved functions and with 50-80% amino acid identity, is a powerful model for the study of pathological mechanisms of disease-related genes and in the screening for potential therapeutics. Zebrafish is a cheapest and fastest alternative to the costly and labor-intensive requirements necessary for the creation of mouse models of neurodegenerative disorders. However, we should be aware that HSP zebrafish models can only reflect at best, a general effect of the gene due to the lack of upper motor neurons in this vertebrate (Kabashi et al., 2011; Patten et al., 2014).

The zebrafish allows the study of most genes either by simple transient assays or by the creation of stable disease models using genome-editing techniques, both by loss- or gain-of-function approaches. The easiest method to study loss-of-function mutations in zebrafish is with the injection of antisense morpholino oligonucleotides, which effectively bind to the target mRNA blocking translation and therefore knocking down expression. This creates transient knockdowns that should be controlled for off-target toxicity with the most important control for disease models being the demonstration of, at least, partial rescue by the homologous human mRNA. However, the temporary nature of this microinjection knockdown could carry limitations when attempting to model late-onset diseases. To generate stable loss-of-function models for neurodegenerative diseases there are some powerful genome-editing techniques like CRISPR/Cas but that can be expensive, complex or also carry some off-target toxicity. mRNA injections could be used to test gain-of-function in a transient way and for stable models, the use of an efficient Tol2 transposon system is a common approach (Kabashi et al., 2011; Patten et al., 2014).

Zebrafish models of HSP to validate the pathogenic nature of genes are becoming increasingly common, with more than 14 genes studied to date, all by transient knockdown. The phenotypes found are usually abnormal motility and motor neuron development defects. Knockdown of *SPAST*, the most frequent cause of HSP, caused deficiencies in the spinal motor axon outgrowth, impaired locomotion and perturbation of axonal microtubule dynamics (Wood et al.,

2006). Knockdown of two recessive HSP genes (SPG11 and SPG15) caused mobility impairment and anomalous branching of spinal cord motor neurons at the neuromuscular junction being suggested, for the first time and *in vivo*, a genetic interaction between these genes (Martin et al., 2012). More recently, Novarino et al used morpholino knockdown in zebrafish for confirmation of the pathogenicity of four identified genes (*ARL6IP1*, *MARS*, *PGAP1* and *USP8*) (Novarino et al., 2014). Up to now no drug test or screens for HSP using zebrafish has been published (Babin et al., 2014; Patten et al., 2014). Drug discovery is many times hampered by the lack of inexpensive models to investigate and to identify candidate drugs, which is even truer in rare diseases like HSP where the low prevalence and large heterogeneity does not motivate commercial companies. However, it is urgent the need to develop effective treatments for HSP as there are still no treatments targeting the underlying pathophysiology.

2. Major goals

2.1 Main aims

This work aimed to genetically characterize a Portuguese cohort of HSP families collected through a national population-based survey that meant to identify the majority of the Portuguese families with hereditary spastic paraplegia. This study started with a homogeneous and large cohort of families without molecular diagnosis (100 families of the 159 with available DNA) already screened for the most frequently mutated HSP genes.

The main goals of this work were:

1) Provide a molecular diagnosis to the Portuguese HSP families,

2) Identify novel genes responsible for hereditary forms of spastic paraplegia,

3) Functionally characterize the identified candidate genes and variants.

2.2 Specific objectives

1. Genetic characterization of the families without molecular diagnosis by panel sequencing

a. Identify causative point variants and large rearrangements in genes already reported

b. Confirm the new molecular diagnosis by segregation analysis

2. Gene identification through linkage analysis and next generation sequencing (NGS)

a. Exome sequencing in some of the largest families to find new HSP-causing genes

b. Linkage analysis to focus the variant search in the candidate regions identified

3. Screen additional families for variants in the new genes identified and establish genotype-phenotype correlations

a. Test the remaining HSP families without molecular diagnosis from the Portuguese and the SPATAX network cohort to explore possible correlations between specific phenotypes and different types of variants

4. Functionally characterize the newly identified genes and candidate variants

a. Bioinformatics analysis to check possible molecular functions and identify potential structural homologies

b. Biological characterization to assess the expression pattern, the sub-cellular localization of the protein and its cellular functions.

3.Material and Methods

3.1. Patients

3.1.1 Inclusion criteria

In Portugal, a population-based survey was conducted from 1994 to 2004 that aimed to identify all the families with hereditary spastic paraplegia and hereditary cerebellar ataxia (Coutinho et al., 2013; Silva et al., 1997). This study was able to identify 412 patients from 193 families with HSP using homogeneous inclusion criteria as the same team of neurologists examined all the patients. For inclusion, patients had to have progressive gait disturbances with corticospinal signs in the lower limbs and preferentially a family history of spastic paraplegia. In this survey it was collected a detailed clinical and familial history, it was performed a neurological examination and, whenever possible, one patient from each family was admitted for complete clinical, serologic and radiologic workup. Written consent for genetic study was requested and DNA from the consenting families was stored (159 families). The clinical data were registered, and all consenting families were genetically tested. The families with an expected autosomal dominant (AD) inheritance mode were tested for disease-causing variants in *ATL1*, *SPAST*, and *REEP1* (Loureiro et al., 2009). In the autosomal recessive (AR) group, patients with cognitive impairment were screened for causative variants in *SPG11*, *ZFYVE26* and *CYP7B1*. These genetic analyses were able to provide a molecular diagnosis to 59 of the 159 families (37,1%).

For the present study, we worked on the 100 remaining families without molecular diagnosis and with an available DNA sample.

3.1.2 Clinical phenotypes

Families were classified into pure (only progressive spastic paraplegia in all affected family members – 56 families) or complex (associated with other neurologic symptoms like neuropathy, ataxia, cognitive impairment, among others - 44 families). The age at onset was defined as the age at which patients or their relatives noticed any walking difficulties, change of gait pattern, cramps or unexplained falls, and were considered as having an early onset when the first symptoms were noticed before the age of 20 years and after that, classified as late onset. Age at onset in the whole cohort ranged from less than 1 year to 65 years. Among these 100 Portuguese families, 44 showed an apparent dominant inheritance, 40 a recessive inheritance and 16 were isolated cases.

	Dominant families	Recessive Families	Isolated cases	Total
Pure Early onset	17	8	3	28
Pure Late onset	18	5	5	28
Complex Early onset	4	21	6	31
Complex Late onset	5	6	2	13
Total	44	40	16	100

3.1.3 Ethics

Written consent for genetic study was obtained from all tested individuals or their respective legal guardian for all the blood samples and skin biopsies. The ethical committees of ICBAS (Portugal) and of the Paris-Necker Hospital (France) approved the genetic analyses.

3.2. Biological material

3.2.1 DNA extraction and quality check

Genomic DNA was extracted from peripheral blood samples through salting-out and later by an automated DNA purification method using the QIASymphony DSP DNA Midi kit (Qiagen, Hilden, Germany).
All DNA samples were quantified and controlled for contamination with proteins or solvents using a NanoDrop® 2000 or an 8000Uv-Vis spectrophotometer (ThermoFisher Scientific, Waltham, USA). Furthermore, DNA quality, was checked by agarose gel electrophoresis to look for the presence of degradation before being used in the NGS experiments and a test PCR was performed to check for PCR inhibitors. An additional quantification using the Qubit® fluorometer (ThermoFisher Scientific, Waltham, USA), which specifically measures double stranded DNA, was performed in the samples before NGS.

3.2.2 Skin biopsies and fibroblast culture

Skin biopsies were requested only when a strong candidate gene was identified by exome analysis, so that patients' fibroblasts could be important to proceed with the functional study. For skin biopsies, local anesthetics were applied to the upper arm and a sterile disposable 2.5mm biopsy punch (Kai medical, Seki, Japan) was used.

The fibroblast culture, as all the other cell cultures were performed under sterile conditions in a laminar flow hood. Skin biopsies were cleaned in a Petri dish with culture medium and cut into small pieces with a scalpel. The small pieces were transferred to two T25 culture flasks with a small drop of heated medium (DMEM (Dulbecco's Modified Eagle Medium) with Glutamax (ThermoFisher Scientific, Waltham, USA), 15% of fetal bovine serum (FBS) (Invitrogen, Carlsbad, USA) and 1% antibiotic-antimycotic (ThermoFisher Scientific, Waltham, USA)) on top of each piece. The flasks were incubated in a humidified CO_2 incubator at 37ºC. Two days later some drops of medium were added to the pieces. After five days 2mL of culture medium were added and after that, the medium was changed every 3 days. The flasks were frequently observed under the microscope and when the fibroblasts were dense and confluent, we did the first trypsinization and transfered the fibroblasts to T75 culture flasks (1º passage).

3.3. Targeted next generation sequencing for screening of known HSP genes

3.3.1 Panel design

A custom next generation sequencing panel that covered 34 HSP-related genes was designed with the NimbleGen SeqCap EZ Choice Library (Roche, Basel, Switzerland) according to the request and approval of the ICM team. This first panel covered 31 known HSP genes responsible for dominant and recessive forms and three genes (*ALS2*, *SACS* and *SETX*) implicated in overlapping phenotypes. The 34 genes corresponded to 531 regions with a total length of 109 768 base pairs.

Later, a second panel was designed by the same method, in order to cover more HSP genes, including some newly identified. This second panel covered 70 genes that added to the previously, 30 new HSP genes, as well as three genes (*FBXO7*, *GJA1* and *SAMHD1*) causing overlapping phenotypes and three candidate genes that encode partners of known HSP proteins (*AP5B1*, *AP5M1* and *AP5S1*). This panel covered a total of 210 363 bases corresponding to 1001 regions. These two panels targeted exonic regions and at least 20 intronic bases at the exon-intron boundaries.

3.3.2 Experimental procedure

The procedure elected consisted of three major steps: library preparation, double capture and massive parallel sequencing. For the library preparation, we used the Illumina TruSeq DNA LT Sample Preparation kit v2-Set A (Illumina, San Diego, USA) according to the manufacturer´s

protocol for the first panel and the Kapa Library Preparation Kit for Illumina platforms (Roche, Basel, Switzerland) with its corresponding protocol for the second panel. First, the samples had to be quantified with the Qubit® fluorometer and diluted to a final volume of 55µL at 20ng/µl. The DNA was then fragmented in a Covaris S220 ultrasonicator (Covaris, Woburn, USA) to fragments between 180-220bp. After fragmentation, we performed end repair to convert the overhangs into blunt ends, adelynation of the 3' ends and adapter ligation to prepare the libraries for multiplexing and hybridization in a flow cell. As a last step before capture there was a pre-capture PCR to enrich the DNA fragments with adapters at both ends and, after cleaning, samples were quantified with the nanodrop and the Caliper® bioanalyzer (PerkinElmer, Waltham, USA) was used to check the library. For the several cleaning steps we used the Agencourt AMPure XP beads (Beckman Coulter, Brea, USA).

In the second part, for the double capture, we used the Roche NimbleGen SeqCap EZ Reagent Kit (Roche, Basel, Switzerland) and the customized SeqCap EZ Library using the SeqCap EZ Library SR User's Guide protocol. We started by multiplexing the samples (12 samples per experiment) by mixing in equal amounts (by mass) each sample to get a combined mass of at least 1,25 µg in a single pool. Then, we did the hybridization of the samples to a half aliquot of SeqCap EZ Library (custom probes), washed and recovered the captured DNA and performed a ligation-mediated PCR. These steps were repeated from the hybridization, now with the other half of aliquot (double capture). At the end of the last cleaning step, samples were quantified with nanodrop and Qubit and the mean size of fragments of the pools was checked on the Caliper bioanalyzer. For the second panel, 2 pools of 12 samples were mixed before sequencing (equal amount by mass).

For the massive parallel sequencing, we used the Illumina MiSeq Benchtop Sequencer (Illumina, San Diego, USA). First, we had to denature and dilute the library and added it to the reagent cartridge. The cartridge, the flow cell, the wash buffer and the empty waste bottle were loaded into the MiSeq instrument. A sample sheet was created for each run with the project information and the correspondence between samples and the indexes used. For the first panel we perform runs of 2* 150bp reads and for the second panel runs of 2* 250bp reads.

3.3.3 Bioinformatic analysis

The analysis of the fastq files produced in each sequence run was done using the CLC Bio-Genomics Workbench 6.5.1 software (https://www.qiagenbioinformatics.com/). In this software, we have created a workflow that performs: (1) the alignment of the reads with the *Homo sapiens* (hg19) reference sequence; (2) a probabilistic and quality-based variant detection; (3) statistics for target regions; and (4) annotations with relevant variant data as amino acid change, predicted splice effect and SNP (single nucleotide polymorphism)

frequencies. Additionally, with these data, we created a local database (updated after every sequencing run) and the variant files were re-annotated with this information.
In order to get the probably pathogenic variants for each patient from the variant files created, several filters were applied. We selected only the variants: (1) present in the target regions, (2) causing a change at the protein level, (3) present in the local database with a frequency below 15% and, (4) if known in the databases, its minor allele frequency (MAF) had to be below 1% for AR inheritance and 0.2% for AD transmission**.** All filtered variants were further analyzed using Alamut v.2.4 software (Interactive Biosoftware, La Rochelle, France) for pathogenicity prediction with SpliceSiteFinder, MaxEntScan, NNSPLICE, GeneSplicer, Human Splicing finder, Polyphen-2, SIFT, MutationTaster, Align GVGD and UMD-Predictor.
Additionally, we looked for the presence of large rearrangements, by the analysis of the coverage data of all regions with an in-house algorithm based on the comparison of the coverage between patients.

3.3.4 Confirmation and segregation analysis

3.3.4.1 Sanger sequencing

We confirmed the presence of the variants identified through NGS by Sanger sequencing. After confirming each variant and when DNA was available we tested also additional family members to check segregation of the variants with the disease. Primers were designed using Primer3Plus software (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi). The PCR amplification was done using the DreamTaq DNA Polymerase Kit (ThermoFisher Scientific, Waltham, USA) or when necessary, AccuPrime GC-Rich DNA polymerase (Invitrogen, Carlsbad, USA), followed by sequencing at GATC BioTech (Konstanz, Germany). Sequence analysis was performed using SeqScape 2.6 software (Applied Biosystems, Foster City, USA).

3.3.4.2 MLPA and quantitative real-time PCR

We confirmed the presence of large rearrangements, suggested by the coverage analysis, by multiplex ligation-dependent probe amplification (MLPA) when available or quantitative real-time PCR (qPCR). In order to confirm the presence of deletions or duplications in *SPG11* we used the SALSA MLPA kit P306 (MRC Holland, Amsterdam, Netherlands) according to manufacturer´s instructions. The resulting fragments were analyzed on an ABI 3130XL Genetic Analyzer using 500-LIZ (Applied Biosystems, Foster City, USA) as a size standard, in the GeneMarker v1.90 software (SoftGenetics, State College, USA). For the genes for which no MLPA kit was available, we performed qPCR. Primers were designed using Beacon Designer (Premier Biosoft, Palo Alto, USA), one in each exon that we needed to verify, and one in each

of the contiguous exons (in the same gene, or in the gene closest to the expected mutated region) as well as one from a control gene present in a different chromosome. A control sample was also amplified for all the amplicons. The PCR amplification was performed using iQ SYBR Green Supermix (Bio-Rad, Hercules, USA) in an iQ5 Real-Time PCR Detection System (Bio-Rad, Hercules, USA). Real-time PCR amplification of each sample and control was done in triplicate. Results were analyzed with IQ5 Optical System Software (Bio-Rad, Hercules, USA) and the heterozygosity was determined by comparison between the control and sample fragments.

3.4. Whole exome sequencing (WES) for identification of new HSP genes

3.4.1 Family selection

We selected the families with the largest number of informative family members with available DNA and without any potential diagnosis found in the NGS panels.

The first families selected for exome were two AR families (SR2 and SR9) not included in the NGS panel study because they had already been tested for several HSP genes without any diagnosis found and in whom whole exome linkage studies were already performed. We selected 2 patients of each family for exome analysis, focused in the locus identified in the whole genome linkage analysis: in family SR2 a new locus for the disease was found in chromosome 14, namely SPG32, (Stevanin et al., 2007b) and, in family SR9, a new locus was identified on chromosome 4 with a significant LOD (logarithm of the odds) score of 4.3.

There were 10 additional families selected for exome sequencing that remained without diagnosis after the NGS panel screening, 5 AD and 5 AR families. In the families with expected dominant transmission, we selected three family members per pedigree taking into consideration their genealogical distance so that they shared the smallest number of variants possible. In the families with presumed recessive inheritance, we selected a parent and an affected child in the families with consanguinity and we focused in the homozygous variants identified and, in the families without consanguinity, we selected a parent and two affected children.

3.4.2 Experimental procedure

For the first four families selected (1 AR and 3 AD), exome sequencing was performed at the Integragen facility (Evry, France). This company used the Agilent SureSelect All Exon kit V5

50Mb (Agilent Technologies, Santa Clara, USA) to capture the exomes and the paired-end sequencing was done with the HiSeq2000 sequencer (Illumina, San Diego, USA).
The other eight families were sequenced at the ICM genetic platform (IgenSeq). The procedure was also based in a capture method (similar to the one described in 3.3.2) but the capture protocols were diverse. The first four families had their exome captured with the Nextera Rapid Capture Exome kit (Illumina, San Diego, USA), two families with the SeqCap EZ Human Exome Library v3.0 of NimbleGen (Roche, Basel, Switzerland), one family with the SureSelect QXT kit (Agilent Technologies, Santa Clara, USA) and the last family with the SureSelect XT kit (Agilent Technologies, Santa Clara, USA). Massive parallel sequencing was performed in the NextSeq500 sequencer (Illumina, San Diego, USA) for all families**.** The procedure was similar to the one applied on the Miseq sequencer, but with the proper cartridges for runs of 2* 75bp reads.

3.4.3 Bioinformatic analysis

The Integragen exomes were analyzed with the Eris platform, whereas the data from the exomes sequenced at ICM were treated by the bioinformatics platform of the Institute. Variants were identified using a custom workflow that used FastQC for the quality control, BWA-0.7.8 for the alignment with the hg19 reference genome, GATK for the coverage analysis and variant calling and ANNOVAR for annotation.
The exomes were then filtered, family by family, taking into consideration the presumed inheritance mode. The variants had to cause a change at the protein level and, if known in the databases, its minor allele frequency (MAF) had to be below 1% for AR inheritance and 0.2% for AD transmission**.** All filtered variants were further analyzed using Alamut v.2.4 software (Interactive Biosoftware, La Rochelle, France) for pathogenicity prediction and the frequency checked on the exac database (http://exac.broadinstitute.org/).

3.4.4 Confirmation and segregation analysis

As described in 3.3.4 we confirmed the presence and checked the segregation of the filtered variants by Sanger sequencing. For these samples, PCR amplification was done using the HotStart Master Mix (Qiagen, Hilden, Germany) or when necessary, GC-Rich PCR System (Roche, Basel, Switzerland), followed by Exo/SAP Go (Grisp, Porto, Portugal) purification and sequencing using BigDye Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, USA) in an ABI-PRISM 3130 XL automatic sequencer (Applied Biosystems, Foster City, USA).

3.5. Genome-wide genotyping and linkage analysis

3.5.1 Family selection

As the exome analysis of the AD families showed that even with three patients tested there were too many variants shared, we decided that it was important to reduce the chromosomal area to focus on. Therefore, we selected these five families with presumed dominant inheritance and exome data for whole genome linkage analysis using SNP microarrays.
Three of the families with expected recessive transmission (SR2, SR9 and SR13) had linkage data available previously performed by the screening of 300 autosomal microsatellites spaced 15cM apart (Stevanin et al., 2007b).

3.5.2 Experimental procedure

We used the Infinium® HumanCore-24 v1.0 DNA Analysis kit (Illumina, San Diego, USA) for the genome-wide genotyping. This BeadChip allows us to genotype 306 670 markers of which 240 000 are highly informative genome-wide tag SNPs. We followed the Infinium High Throughput Screening automated assay protocol from Illumina. After quantification, the DNA samples were denatured and neutralized to prepare them for an overnight incubation at 37ºC in a hybridization oven, that leads to a uniform amplification of the genomic DNA. In the next day, the DNA samples were enzymatically fragmented, precipitated, resuspended in hybridization buffer and hybridized to the BeadChip overnight. After washing the BeadChip to remove all non-specifically hybridized DNA, there is a single-base extension and staining on the BeadChip responsible for incorporating labeled nucleotides in the extended primers using the captured DNA as template. Lastly, the beadchip was imaged with the compatible iScan system (Illumina, San Diego, USA) that records the images resulting from the excited fluorophore inserted in the single-base extension. The system provides the intensity data files (.idat) for analysis.

3.5.3 Bioinformatic analysis

The GenomeStudio V2011.1 software (Genotyping Module) (Illumina, San Diego, USA) was used for data analysis. A genotyping quality check was first performed and the genotypes were determined for all markers. After removing all the uninformative and poor-quality markers, the Merlin Input Report Plug-in v2.0.1 was used to create the input files for the Merlin application. The pedigree files were completed, and the linkage analysis was done with the Merlin application v1.1.2 (Abecasis group, University of Michigan USA) following the model for autosomal dominant disorders.

3.6. Functional studies

3.6.1 DNA constructs and site-directed mutagenesis

To study the new HSP candidate genes, we used the Myc-DDK-tagged-Human AKAP6 (Origene, Rockville, USA) encoding the full length AKAP6 protein. We cloned also the total cDNA of *AIMP1* (NM_001142416.1) into the pEGFP-C1 vector (Clontech, Mountain View, USA). To construct the *AIMP1* minigene we cloned the exons 2 to 4 (with introns) from the patient sample into the pCMVdi vector kindly provided by Dr. Alexandra Moreira (Glória et al., 2014). To create the mutant constructs, we performed site-directed mutagenesis using the QuikChange Site-Directed Mutagenesis Kit (Agilent Technologies, Santa Clara, USA) according to the manufacturer's protocol with the appropriate primer pairs designed with the QuikChange Primer Design Program (Agilent Technologies, Santa Clara, USA).

3.6.2 Cell culture and transfection (HEK293T and fibroblasts)

The medium used in the maintenance of HEK293T and human fibroblasts was DMEM-Glutamax (ThermoFisher Scientific, Waltham, USA) with 10% FBS (Invitrogen, Carlsbad, USA) and 1% antibiotic-antimycotic (ThermoFisher Scientific, Waltham, USA). Cells were cultured on 25 or 75 cm^2 culture flasks in a humidified 5% CO_2 atmosphere at 37ºC. Upon confluency, cells were trypsinised and passed to new flasks. Cells were frozen with 10% DMSO (Sigma-Aldrich, St. Louis, USA).

Transfection of fibroblasts was done with FuGENE HD (Roche, Basel, Switzerland) and for the HEK293T cells we used the jetPrime (Polyplus, Berkeley, USA) transfection reagent (24h after being plated). Cells were plated in the adequate concentration 24h before transfection. The transfection mixture and incubation time was performed according to the manufacturer's protocol.

3.6.3 Antibodies and subcellular fractionation

The primary antibodies used were: rabbit polyclonal anti-mAKAP (Covance, Princeton, USA), mouse monoclonal anti-GM130, anti-TOM20 (BD Biosciences, Franklin Lakes, USA) and anti-α-tubulin clone B-5-1-2 (Sigma-Aldrich, St. Louis, USA) to mark the Golgi, mitochondria and microtubules respectively; rabbit polyclonal anti-LAMP1 (Abcam, Cambridge, UK) and anti-calreticulin (Enzo Life Sciences, Farmingdale, USA) to mark the lysosome and the ER respectively; also the monoclonal mouse and rabbit Myc-tag antibodies (Cell Signaling Technology, Danvers, USA) that recognizes Myc-fusion proteins.

The secondary antibodies used for western blot were goat anti-mouse IgG-HRP (Santa Cruz Biotechnology) and goat anti-rabbit IgG HRP-conjugate (Millipore, Billerica, USA) and for immunostaining we used the goat anti-mouse IgG, Alexa Fluor 488, 568 or 647 conjugate and the goat anti-rabbit IgG, Alexa Fluor 488 or 568 conjugate (Thermo Fisher Scientific, Waltham, USA).
For subcellular protein fractionation, we used the Subcellular Protein Fractionation Kit for Cultured Cells (ThermoFisher Scientific, Waltham, USA) according to the manufacturer's protocol.

3.6.4 Western blot

For protein expression analysis, cells were collected and lysed on ice with RIPA buffer (Sigma-Aldrich, St. Louis, USA) with protease inhibitors (cOmplete™, EDTA-free Protease Inhibitor Cocktail – Roche, Basel, Switzerland). Samples were then sonicated and vortexed. Protein concentration was determined using the Pierce BCA Protein Assay Kit (Thermo Fisher Scientific, Waltham, USA).
Lysates were re-suspended in Laemmli loading buffer with β-mercaptoethanol (Sigma-Aldrich, St. Louis, USA) and denatured at 95ºC for 5 minutes. 20µg of each sample was loaded on a gradient gel and then transferred to a polyvinylidene fluoride (PVDF) membrane (Millipore, Billerica, USA). After blotting, PVDF-membranes were blocked with 3% nonfat dry milk in PBS containing 1% Tween-20 (PBS-T) for 1 h with agitation. Membranes were probed with the selected antibodies and after washing, the proteins were detected with WesternBright Sirius HRP substrate (Advansta, Menlo Park, USA).

3.6.5 Immunofluorescence

Cells were grown on glass coverslips washed and coated with fibronectin (Sigma-Aldrich, St. Louis, USA). After 48h, cells were fixed in 4% paraformaldehyde/sucrose in PBS for 20 minutes at room temperature and permeabilization was done with 0.3% Triton X-100 or, for lysosome imaging, with microscopy buffer (Harris et al., 2013). Nuclei were stained with Hoechst 33258 (Invitrogen, Carlsbad, USA) and cells were mounted with Prolong Gold (Invitrogen, Carlsbad, USA) and sealed with nail polish. Images were obtained on a Zeiss Axio Imager Z1 (Zeiss, Oberkochen, Germany) with a coupled device camera, using 20X or 100X (oil) objectives. Images were acquired in the AxioVision release 4.8 and analyzed with the open-source software CellProfiler (Carpenter et al., 2006).

3.6.6 RNA extraction, RT-PCR and qPCR

RNA was extracted from HEK293T cells at 48h after transfection and in the fibroblasts cells, when confluent in 100mm culture dishes. Extraction was performed by a double method with NZYol reagent (Nzytech, Lisboa, Portugal) and the RNeasy Mini Kit (QIAGEN, Hilden, Germany). After nanodrop quantification, cDNA was synthesized by reverse transcription using SuperScript III First-Strand Synthesis System (Invitrogen, Carlsbad, USA) according to manufacturer's instructions.

Primers for the qPCR were designed using Beacon Designer (Premier Biosoft, Palo Alto, USA), selecting primer pairs in different exons to assure that genomic DNA was not amplified. PCR amplification was performed using PowerUp SYBR Green Master Mix (Thermo Fisher Scientific, Waltham, USA) in a 7500 Fast Real-Time PCR (Thermo Fisher Scientific, Waltham, USA). Real-time PCR amplification of each sample and control (we used *ACTB* levels) was done in triplicate. Results were analyzed with 7500 Software v2.0.6 (Thermo Fisher Scientific, Waltham, USA).

3.6.7 Zebrafish studies

Transgenic embryos (Hb9: GFP line) that specifically express GFP (Green Fluorescent Protein) in spinal motor neurons were used (Flanagan-Steet et al., 2005). For in situ hybridization, two sense and antisense RNA probes of the *AKAP6* transcript were synthesized using T7 and T3 Polymerases (Promega). An antisense or morpholino (MO) oligonucleotide directed against the translation initiation site of *AKAP6* was designed by GeneTools (5' TTGGTGATATGACAACGTTCATGGT 3') that was injected at the 2-cell stage with a microinjector (Eppendorf). For escape experiments in response to touch, a tactile stimulus was applied to the larvae at 72 hpf. Image acquisition was performed using a LEICA M165C binocular loupe equipped with an IC80HD camera. Response was recorded with a LEICA M165C binocular loupe equipped with an IC80HD camera. Immunohistochemistry was performed in the embryos by fixation in 4% paraformaldehyde (PFA) and permeabilization with trypsin (0.25%). Embryos were subsequently incubated overnight with a specific primary antibody and the next day with the corresponding secondary antibody. The antibodies used were: mouse anti-znp1 (1:100) and anti-zn5 (1:200) (ZIRC) to mark the primary and secondary motor neurons respectively, mouse anti-F59 (1:100) (Santa Cruz) to mark the slow muscle fibers and Rabbit anti-GFP (1:1000) (Invitrogen) to mark GFP. The images were acquired using a fluorescence microscope (Zeiss, Axiovert 200M) equipped with an Apotome and analyzed with ImageJ software (NIH). Each image corresponds to a planar projection (z-stack) of 10 to 20 images of 2 µm thickness.

3.7. Targeted next generation sequencing for screening of new candidate HSP genes (Fluidigm assay)

3.7.1 Panel design and family selection

In order to study some of the new candidate genes identified in the laboratory, a multiplex assay for next generation sequencing was designed that covered 19 genes. It included twelve candidate genes (including *AKAP6*) and seven recently published HSP or HSP-related genes: *ALDH18A1*, *CAPN1*, *DNM2*, *FARS2*, *GCH1*, *KLC4* and *TUBB4A*.

The families selected were the ones where no molecular diagnosis was found in the first strategy (NGS panels) that included 67 Portuguese HSP families and 125 families from the SPATAX cohort.

3.7.2 Experimental protocol

For the study of this mini-panel we opted for an amplicon-based library preparation followed by NGS. The preparation of the library was performed using the Fluidigm Access Array™ IFC 2-Primer (Fluidigm, South San Francisco, USA) that allows the simultaneous enrichment of 48 to 480 amplicons (if multiplex) per sample from 48 samples.

Primers were designed using Primer3Plus software and the universal sequences Forward (AAGACTCGGCAGCATCTCCA) and Reverse (GCGATCGTCACTGTTCTCCA) were added to be used as adapters. The primers for the 302 fragments were then PCR tested using DreamTaq DNA Polymerase Kit (Thermo Fisher Scientific, Waltham, USA) or, when necessary (usually for the first exons), AccuPrime GC-Rich DNA polymerase (Invitrogen, Carlsbad, USA) and checked on the Caliper bioanalyzer (PerkinElmer, Waltham, USA). After confirmation, the 277 primers pairs (excluding first exons) were pooled in 144 wells for multiplex PCR. To perform the PCR, first we prepared the Access Array plate and loaded the primers dilutions, PCR mix and DNA samples (48 of each/run) in the IFC controller AX machine. The PCR was then performed in the Fluidigm FC1 cycler. To recover the amplicons, we used another IFC controller AX (Post PCR). All the PCRs of the same samples were then mixed (192 libraries) and a second PCR was performed. This is essential for adding the sequencing adapters P5 and P7, recognized by the Illumina flowcell, which contains a barcode that will allow the sample identification. After confirmation on Caliper and purification, a nanodrop quantification was performed and all the 192 libraries were mixed in one pool (equal amount by mass). After purification of the pool, verification on the caliper bioanalyzer and QuantiFluor quantification (Promega, Madison, USA), samples were sequenced in the Illumina MiSeq Benchtop Sequencer (Illumina, San Diego, USA) with the MiSeq Reagent Kit v3 (600-cycle). Library and MiSeq preparation as described in 3.3.2.

There was a second run of sequencing performed on the MiSeq, as described, which included the first exons. The library preparation was slightly different and the 25 amplicons were PCR amplified in singleplex using the Bravo Automated Liquid Handling Platform (Agilent Technologies, Santa Clara, USA). After amplification, all the PCRs from the same samples were pooled together, and, as described for the other amplicons, indexed with the P5 and P7 sequencing adapters and then, purified, quantified and pooled in a unique library for massive sequencing.

3.7.3 Bioinformatic analysis

The analysis of the fastq files produced in each sequence run was performed by the bioinformatics platform at ICM as described in 3.4.3. As before, the variants had to be present in the target regions, cause a change at the protein level and, if known in the databases, its MAF had to be below 1% for AR inheritance and 0.2% for AD transmission. All filtered variants were further analyzed using Alamut v.2.4 software (Interactive Biosoftware, La Rochelle, France) for pathogenicity prediction.

3.7.4 Confirmation and segregation analysis and *AIMP1* study

We confirmed the presence of the variants identified in the mini panel by Sanger sequencing of the region where the variants were present (as described in 3.4.4).
In addition, due to the evidence found that pointed out *AIMP1* as a strong candidate gene, we performed Sanger sequencing of the seven exons of this gene in the 67 Portuguese HSP families without molecular diagnosis. Primers were designed for the entire gene with the Primer3Plus software and the PCR amplification, purification and sequencing was performed as already described.

4. Results

4.1 Search for molecular diagnosis in already identified genes

The study of a large Portuguese cohort of 193 families with hereditary spastic paraplegia for the most frequently mutated genes: *ATL1*, *SPAST*, and *REEP1* (Loureiro et al., 2009) in AD families and *SPG11*, *ZFYVE26* and *CYP7B1* in AR families achieved a percentage of molecular diagnosis in 37,1% of the families (59 of the 159 families with available DNA). In order to provide a diagnosis to the remaining families, we started by screening 98 of these families for variants in genes already associated with HSPs. Two custom panels for targeted next generation sequencing were used, which screened 34 genes in the first panel and 70 genes with the second panel. This allowed us to provide a molecular diagnosis to 20 additional families and moreover, we identified variants of unknown significance in another 37 families that could led to possibly even more diagnosis if additional family members of these families could be tested or if functional studies are performed, which could help to prove the pathogenicity of the variants identified.

This work was published at the European Journal of Human Genetics:

Article 1: Morais S, Raymond L, Mairey M, Coutinho P, Brandão E, Ribeiro P, Loureiro JL, Sequeiros J, Brice A, Alonso I, Stevanin G: “Massive Sequencing of 70 Genes Reveals a Myriad of Missing Genes or Mechanisms to be uncovered in Hereditary Spastic Paraplegia”. European Journal of Human Genetics 2017 Nov;25(11):1217-1228. doi: 10.1038/ejhg.2017.124

EJHG Open

European Journal of Human Genetics (2017), 1–12
Official journal of The European Society of Human Genetics
www.nature.com/ejhg

ARTICLE

Massive sequencing of 70 genes reveals a myriad of missing genes or mechanisms to be uncovered in hereditary spastic paraplegias

Sara Morais[1,2,3,4,5,6,7,8], Laure Raymond[4,5,6,7,8], Mathilde Mairey[4,5,6,7,8], Paula Coutinho[1,2], Eva Brandão[9], Paula Ribeiro[9], José Leal Loureiro[1,2,9], Jorge Sequeiros[1,2,3], Alexis Brice[4,5,6,7,10], Isabel Alonso*[,1,2,3,11] and Giovanni Stevanin[4,5,6,7,8,10,11]

Hereditary spastic paraplegias (HSP) are neurodegenerative disorders characterized by lower limb spasticity and weakness that can be complicated by other neurological or non-neurological signs. Despite a high genetic heterogeneity (>60 causative genes), 40–70% of the families remain without a molecular diagnosis. Analysis of one of the pioneer cohorts of 193 HSP families generated in the early 1990s in Portugal highlighted that *SPAST* and *SPG11* are the most frequent diagnoses. We have now explored 98 unsolved families from this series using custom next generation sequencing panels analyzing up to 70 candidate HSP genes. We identified the likely disease-causing variant in 20 of the 98 families with *KIF5A* being the most frequently mutated gene. We also found 52 variants of unknown significance (VUS) in 38% of the cases. These new diagnoses resulted in 42% of solved cases in the full Portuguese cohort (81/193). Segregation of the variants was not always compatible with the presumed inheritance, indicating that the analysis of all HSP genes regardless of the inheritance mode can help to explain some cases. Our results show that there is still a large set of unknown genes responsible for HSP and most likely novel mechanisms or inheritance modes leading to the disease to be uncovered, but this will require international collaborative efforts, particularly for the analysis of VUS.

European Journal of Human Genetics advance online publication, 23 August 2017; doi:10.1038/ejhg.2017.124

INTRODUCTION

Hereditary spastic paraplegias (HSP) are a clinically and genetically heterogeneous group of rare neurodegenerative disorders. HSP have a variable age at onset and are mainly characterized by progressive lower limb spasticity and weakness. These hallmarks are caused by degeneration of the motor axons that mostly affects the distal ends of the long central nervous system tracts.[1] Their estimated prevalence ranges from 1–10/100 000 depending on the geographic localization[2] and is 4.1/100 000 in Portugal.[3]

These disorders are historically divided on clinical grounds into pure and complex forms, according to the absence or presence of additional neurological and extraneurological features, such as neuropathy, cognitive impairment, cerebellar ataxia and epilepsy. Consequently, ancillary tests such as brain/spinal cord magnetic resonance imaging, electroneuromyography and others can show various abnormal patterns. Genetically, all classical modes of inheritance have been described, with 79 SPG-associated loci and more than 60 identified genes, a number that is constantly increasing.[4–10] Moreover, at least 11 additional genes have also been associated with spastic paraplegia, some of them very recently.[11,12] Although a considerable number of genes have already been identified, most of the families worldwide have not been extensively screened so far and remain without a molecular diagnosis, as was the case of 68.9% families in Portugal, before this study. Indeed, among the HSP a considerable number of genes and variants are apparently orphan, affecting single families.

Identification of the responsible genes has a great impact on patients and their families as it reveals the cause of the disease, and may also be important to better understand how these variants arise (founder effects and susceptibility chromosomes, among others) and how they lead to specific phenotypes. This study aimed to provide a molecular diagnosis for the Portuguese families identified through a population-based survey[3] using targeted next generation sequencing (NGS) to screen these patients for the presence of variants in 70 genes already associated with HSP or candidate genes, covering all classical modes of inheritance.

MATERIALS AND METHODS

Patients

In Portugal, 193 families with hereditary spastic paraplegia were identified in a national, population-based survey.[3,13] After exclusion of the most common genes in most of the families and isolated cases (*ATL1*, *SPAST* and *REEP1* in families with dominant inheritance, and *SPG11*, *ZFYVE26* and *CYP7B1* in the recessive ones), there were still 68.9% (133/193) of the families without a

[1]UnIGENe, Instituto de Biologia Molecular e Celular (IBMC), Universidade do Porto, Porto, Portugal; [2]Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Porto, Portugal; [3]Instituto de Ciências Biomédicas de Abel Salazar (ICBAS), Universidade do Porto, Porto, Portugal; [4]INSERM, U 1127, Paris, France; [5]CNRS, UMR 7225, Paris, France; [6]Sorbonne Universités, UPMC Univ Paris 06, UMRS_1127, Paris, France; [7]Institut du Cerveau et de la Moelle épinière, ICM, Paris, France; [8]Ecole Pratique des Hautes Etudes, EPHE, PSL Research University, Paris, France; [9]Serviço de Neurologia, Centro Hospitalar de Entre o Douro e Vouga, Santa Maria da Feira, Portugal; [10]APHP, Hôpital de la Pitié-Salpêtrière, Département de Génétique, Paris, France
*Correspondence: Dr I Alonso, UnIGENe, Instituto de Biologia Molecular e Celular (IBMC), Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Rua Alfredo Allen, 208, Porto 4200-135, Portugal. Tel: +351 220 408 800; E-mail: ialonso@ibmc.up.pt
[11]These authors contributed equally to this work.
Received 23 November 2016; revised 9 June 2017; accepted 27 June 2017

molecular diagnosis, 62% (98/159) if we consider only the families with DNA available for testing. Taking that into consideration, 98 unsolved families (ie, with DNA available) were selected to be screened for variants in known genes using one patient per family.

Among these 98 Portuguese families, 38 showed a probable recessive inheritance, 44 a dominant inheritance and 16 were isolated cases. Pure and complex forms were present and age at onset in the whole cohort ranged from less than 1 year to 65 years. Written informed consent for genetic testing was obtained from all tested individuals or their respective legal guardian and the genetic analyses were approved by the ethical committees of ICBAS (Portugal) and of the Paris-Necker Hospital (France).

HSP panels

Two sequential custom NGS panels were developed at ICM, France, to search for variants in HSP-related genes (Supplementary Table S1). This strategy was favored to exome sequencing for ethical concerns regarding the risk of secondary or incidental findings in HSP diagnosis in agreement with our local ethics committees and for the cost effectiveness given the number of genes involved in HSP that could carry a mutation. The first panel covered 34 genes among which 31 were known to be responsible for dominant and recessive HSP forms and the other three (*ALS2*, *SACS* and *SETX*) were implicated in overlapping phenotypes.[14–16] The 34 genes corresponded to 531 regions with a total length of 109 768 base pairs. In order to increase achieving a molecular diagnosis, a second panel was later designed that covered 70 genes. This added 30 newly identified genes[5] to the previous panel as well as three genes (*FBXO7*, *GJA1* and *SAMHD1*) causing overlapping phenotypes[17–19] and three candidate genes that encode partners of known HSP proteins (*AP5B1*, *AP5M1* and *AP5S1*). This panel covered a total length of 210 363 bases corresponding to 1001 regions. These two panels targeted exonic regions and at least 20 intronic bases at the exon-intron boundaries.

Genetic analysis was performed using a customized Roche/Nimblegen capture followed by NGS in the MiSeq apparatus (Illumina, San Diego, CA, USA). The procedure consisted of four major steps: (1) library preparation, where we used the Illumina TruSeq DNA LT Sample Preparation kit v2-Set A (Illumina) according to the manufacturer's protocol for the first panel and the Kapa HTP Library Prep Kit Illumina (Roche, Basel, Switzerland) with its corresponding protocol for the second panel; (2) double capture using the Roche NimbleGen SeqCap EZ Reagent Kit and SeqCap EZ Library (Roche) using the SeqCap EZ Library SR User's Guide protocol; (3) massive parallel sequencing using the Illumina MiSeq Benchtop Sequencer; and (4) data analysis using the CLC Bio-Genomics Workbench 6.5.1 software (https://www.qiagen-bioinformatics.com/). After the alignment with the *Homo sapiens* (hg19) reference sequence, data were filtered taking into consideration that the variants had to: (i) be present in the target regions, (ii) cause a change at the protein level (nonsense, amino acid change or a splicing effect), (iii) be present in the local database with a frequency below 15% and, (iv) if known in the databases, its minor allele frequency had to be below 1% for autosomal recessive inheritance and 0.2% for autosomal dominant transmission. All filtered variants were further analyzed using Alamut v.2.9.0 software (Interactive Biosoftware, La Rochelle, France) for functional effect prediction with SpliceSiteFinder, MaxEntScan, NNSPLICE, GeneSplicer, Human Splicing finder, Polyphen-2, SIFT, MutationTaster, Align GVGD and UMD-Predictor.

Twelve patients per experiment were sequenced using the first panel, while 24 were analyzed simultaneously using the second panel. The first 25 Portuguese patients tested were sequenced using the first panel and the remaining 73 patients were screened with the second panel.

In order to detect rearrangements, we analyzed the coverage of all the regions with an in-house algorithm based on the analysis of the coverage. In the genes where a suggestive alteration in dosage was detected, multiplex ligation-dependent probe amplification or quantitative real-time PCR was applied to confirm the deletion/duplication.

All the variants that were not excluded after filtering and segregation analysis (whenever possible) were submitted to ClinVar (http://www.ncbi.nlm.nih.gov/clinvar/). ClinVar accession numbers are SCV000574442–SCV000574511.

Sanger sequencing

We confirmed the presence of the variants identified by NGS checking, by Sanger sequencing, the region where the variant was present, including additional family members when DNA was available. Primers were designed using Primer3Plus software (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi). The PCR amplification was done using the DreamTaq DNA Polymerase Kit (ThermoFisher Scientific, Waltham, MA, USA) or when necessary, AccuPrime GC-Rich DNA polymerase (Invitrogen, Carlsbad, CA, MA, USA), followed by sequencing by GATC BioTech using BigDye chemistry in an ABI3730 sequencer (Applied Biosystems, Foster City, CA, USA). Sequence analysis was performed using SeqScape 2.6 software (Applied Biosystems).

Multiplex ligation-dependent probe amplification

In order to confirm the presence of large gene rearrangements in the *SPG11* gene we performed multiplex ligation-dependent probe amplification using the SALSA MLPA kit P306 (MRC Holland, Amsterdam, Netherlands) according to manufacturer's instructions and analyzed the resulting fragments on an ABI 3130xl Genetic Analyzer using 500-LIZ (Applied Biosystems), as a size standard, and GeneMarker v1.90 (SoftGenetics, State College, PA, USA).

RESULTS

Study design

A total of 98 index cases of Portuguese families without molecular diagnosis were screened for variants using either the first or the second custom sequencing panel. The mean coverage varied from 126 to 801 reads: 419 reads per base with 94% of the regions with a coverage $\geq 30\times$ with the first panel and 362 reads with 97% of the regions with at least $30\times$ of coverage with the second panel. All the variants considered relevant (rare, predicted deleterious, expected to have an impact given the suspected mode of inheritance) were confirmed by Sanger sequencing and segregation analysis was performed in all the available family members. The regions with less than 30-fold coverage were screened by Sanger sequencing in autosomal recessive cases in which only one heterozygous potentially disease-causing variant was identified. Relevant insufficiently covered exons in patients with a compatible phenotype were also sequenced by Sanger (ie, exon 1 of *SPG7* in patients with cerebellar atrophy and/or optic atrophy). An in-house index of coverage was calculated for each exon and for each patient to search for genomic rearrangements.

Genetic characterization

A molecular diagnosis of HSP was confirmed in 20.4% (20/98) of the cases. The coverage analysis was responsible for the identification of one large homozygous deletion encompassing exons 12 to 14 in *SPG11* that was confirmed by multiplex ligation-dependent probe amplification (Table 1, Supplementary Figure S1) and was considered to affect spatacsin function. Additionally, after filtering of the NGS data, we found 139 nucleotide variants present in 74 families with minor allele frequency below 1.5% that potentially affected protein function according to at least one prediction software. These variants were divided in: (i) Disease-causing variants (17 variants in 14 families), (ii) Likely disease-causing variants (6 variants in 5 families), (iii) variants of unknown significance – VUS (52 variants in 37 families) and (iv) Excluded variants (64 variants in 45 families) by several criteria detailed in Figure 1. In the remaining 24 families, no disease-causing variants were found after filtering.

We considered 24 (of the 140) variants to be affecting protein function in 20 families (Table 1, Figure 2 and Supplementary Figure S1). Most ($n=18$) were classified as disease-causing variants because they were already reported as disease-causing variants, caused an early stop codon (nonsense and frameshift) or were large gene rearrangements. The other ones ($n=6$) were considered as likely

Table 1 Disease-causing and likely disease-causing variants found in 20 families

Family code	Inheritance mode[a]	Mutated gene	Accession.Version numbers	Exon[b]	cDNA	Predicted consequence	State	N° software [c]	Segregation tested	Reference	ClinVar accession number
(a) Disease-causing variants											
SR88	Recessive	*C19orf12*	NM_001031726.2	E2	c.157G>A	p.(Gly53Arg)	Hom	3/3	No	[20] in NBIA	SCV000574442
SR97	Recessive	*CYP7B1*	NM_004820.4	E4	c.914dup	p.(Met305Ilefs*16)	Het	frameshift	Yes	This study	SCV000574443
				E3	c. 825T>A	p.(Tyr275*)	Het	nonsense	Yes	[21]	SCV000574444
S93	Recessive	*CYP7B1*		E6	c.1456C>T	p.(Arg486Cys)	Hom	3/5	Yes	[21]	SCV000574445
SR84	Recessive	*GBA2*	NM_020944.2	E17	c.2618G>A	p.(Arg873His)	Het	4/5	Yes	[22]	SCV000574446
				E12	c.1888C>T	p.(Arg630Trp)	Het	4/5	Yes	[23]	SCV000574447
SR75	Recessive	*SPG7*	NM_003119.2	E11	c.1454_1462del	p.(Arg485_Glu487del)	Hom	Inframe deletion	Yes	[24]	SCV000574448
SR16	Recessive	*SPG11*	NM_025137.3	E4	c.733_734del	p.(Met245Valfs*2)	Hom	frameshift	Yes	[25]	SCV000574449
SR82	Recessive	*ZFYVE26*	NM_015346.3	E34	c.6296dup	p.(Asn2100Glufs*12)	Hom	frameshift	Yes	[26]	SCV000574450
S31	Dominant	*KIF5A*	NM_004984.2	E8	c.610C>T	p.(Arg204Trp)	Het	4/5	Yes	[27]	SCV000574451
S41	Dominant	*KIF5A*		E10	c.839G>A	p.(Arg280His)	Het	5/5	Yes	[28]	SCV000574452
S61	Dominant	*NIPA1*	NM_144599.4	E4	c.316G>A	p.(Gly106Arg)	Het	4/5	Yes	[29]	SCV000574453
S26	Classified as dominant	*SPG7*	NM_003119.2	E10	c.1447C>T	p.(Gln483*)	Hom	nonsense	No	This study	SCV000574454
SR98	Isolated case	*KIF1A*	NM_001244008.1	E2	c.31C>T	p.(Arg11Trp)	Het *de novo*	4/5	Yes	This study	SCV000574455
SR4	Isolated case	*SPAST*	NM_014946.3	E8	c.1133T>A	p.(Leu378Gln)	Het *de novo*	4/5	Yes	[30]	SCV000574456
CI3	Isolated case	*SPG7*	NM_003119.2	E11	c. 1529C>T	p.(Ala510Val)	Het	5/5	No	[31]	SCV000574457
				E3	c.376G>C	p.(Glu126Gln)	Het	3/5	No	[32]	SCV000574458
SR17	Isolated case	*SPG11*	NG_008885.1 (NM_025137.3)	E12-14	c.(2244+1_2245-1)_(2620+1_2621-1)del	r.?; p.?	Hom	deletion	Yes	This study	SCV000574459
(b) Likely disease-causing variants											
S75	Recessive	*ERLIN2*	NM_007175.6	E12	c.899A>T	p.(Asp300Val)	Hom	5/5	Yes	This study	SCV000574460
SR6	Recessive	*SACS*	NM_014363.5	E8	c.[1640C>T; 1634G>T]	p.([Pro547Leu; Trp545Leu])	Hom (both)	2/5; 4/5	Yes	This study	SCV000574461
S27	Dominant	*KIF5A*	NM_004984.2	E10	c.833C>T	p.(Pro278Leu)	Het	5/5	Yes	This study	SCV000574462
S58	Dominant	*KIF5A*		E10	c.868G>C	p.(Asp290His)	Het	4/5	Yes	This study	SCV000574463
SR99	Isolated case	*REEP1*	NM_001164730.1	E3	c.79G>C	p.(Ala27Pro)	Het	4/5	Yes	This study	SCV000574464

[a]The inheritance mode is the one assumed by pedigree analysis.
[b]Exon Number according to Alamut software v2.9.0 that used the NCBI RefSeq numbering scheme (along with simple systematic numbering) on the GRCh37 (hg19) assembly.
[c]Number of Software predicting a functional effect of all the prediction software for the missense variants; Het, Heterozygous; Hom, Homozygous; NBIA, Neurodegeneration with Brain Iron Accumulation.

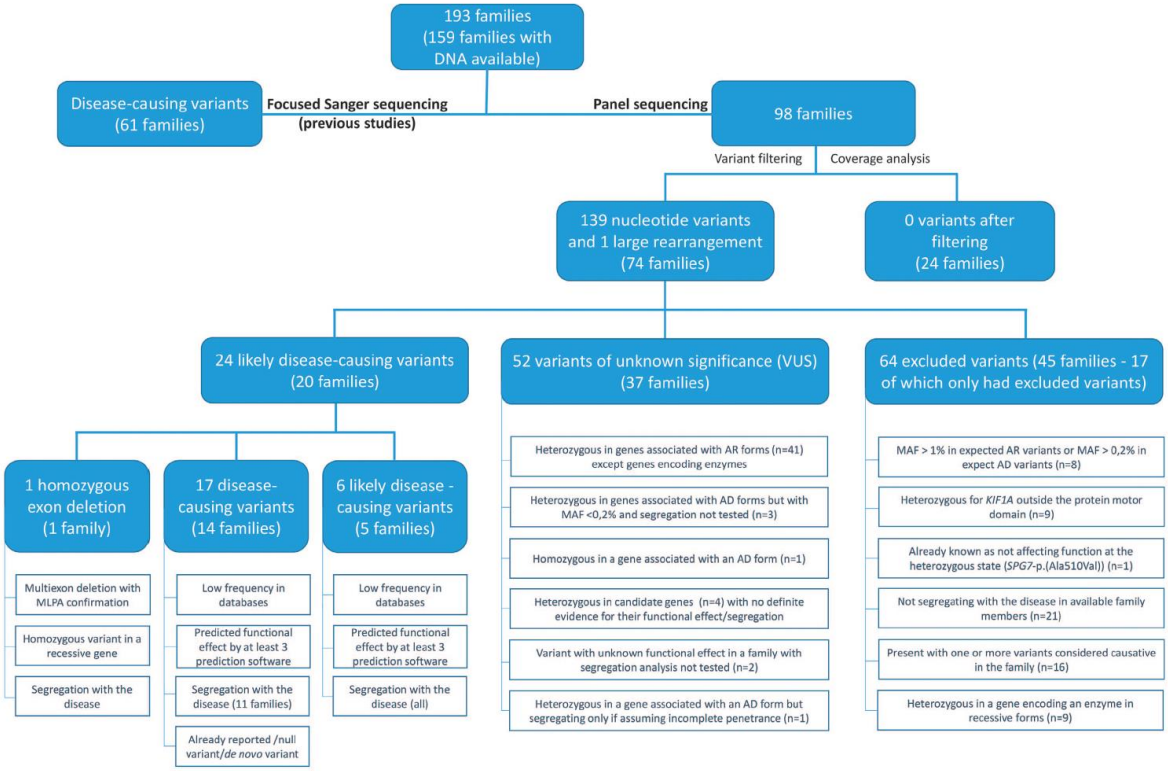


Figure 1 Schematic illustration of the Portuguese cohort with the criteria applied for filtering and classification of the variants identified in the 98 families.

disease-causing because we did not perform functional studies to prove it. All of them had low frequency in databases and a predicted deleterious effect (Table 1) and additionally, in 16 of these families, we could demonstrate their segregation with the disease (Figure 2 and Supplementary Figure S1), which included all the families with likely disease-causing variants.

Among the 52 VUS (Table 2), three are illustrative of the difficulties we faced in the interpretation of their biological relevance (Supplementary Figure S2): a homozygous missense variant in a gene causing a recessive form where family members were not available for segregation analysis (family SR26); a variant in a gene causing a dominant form, of interest only if we consider incomplete penetrance (family SR49) and a variant in the homozygous state in a gene causing a dominant form that segregated with the disease (family S88). In addition, in families SR65 and S86 we found compound heterozygous variants in two different genes known to interact at the protein level (Supplementary Figure S2). In both cases, the variants were found in two different members or interactors of the adaptor protein complex 5, but neither their cosegregation nor their effect at the cellular level could be analyzed, which would represent a prerequisite to prove these potential cases of digenism.

In this study, we report 10 new disease-causing variants among the 24 variants affecting protein function, of which seven are new missense variants (Figure 2), one is a nonsense variant, one a frameshift variant and the last is the deletion of three consecutive exons in *SPG11* (Table 1).

The yield of mutational results varied according to inheritance mode (Figure 3). We found the disease-causing variants in nine of the 38 recessive families (23.7%) and in six of the 44 families with dominant transmission (13.6%). Of note, one family with a homozygous *SPG7* variant was initially considered as dominant (Family S26, Supplementary Figure S1) because of a suspected family history of the disease in the father, who however was not examined and not tested. Five of the 16 isolated cases were also explained at the molecular level (31.3%), including one confirmed *de novo* variant in *KIF1A* (Family SR98, Figure 2). Segregation analysis could not be performed in three families, but all of them had variants already described as disease-causing in the literature.

Clinical characterization

The phenotype of the patients was mostly similar to that of previously reported series. A detailed phenotypic description of the families is presented in Table 3. For example, in the two families carrying *SPG11* variants, the clinical picture included early age at onset and a complex phenotype with cognitive impairment in both families and neuropathy in one of them. Magnetic resonance imaging was not available to check for the presence of a thin corpus callosum. In another family with a homozygous *ZFYVE26* variant and an SPG11-like phenotype, thin corpus callosum was overt. Similarly to published cases, a pure form of the disease was observed in (i) the case with a *de novo SPAST* variant (SR4, Supplementary Figure S1), (ii) the patient presenting a variant in *REEP1* (SR99), (iii) the two families with *CYP7B1* variants (families S93 and SR97, Supplementary Figure S1) and the patient with an *NIPA1* variant (S61, Supplementary Figure S1). On the contrary, in our *KIF5A* families (S41, S31, S27, S58, Figure 2 and Supplementary Figure S1), pure and complex forms were found, in

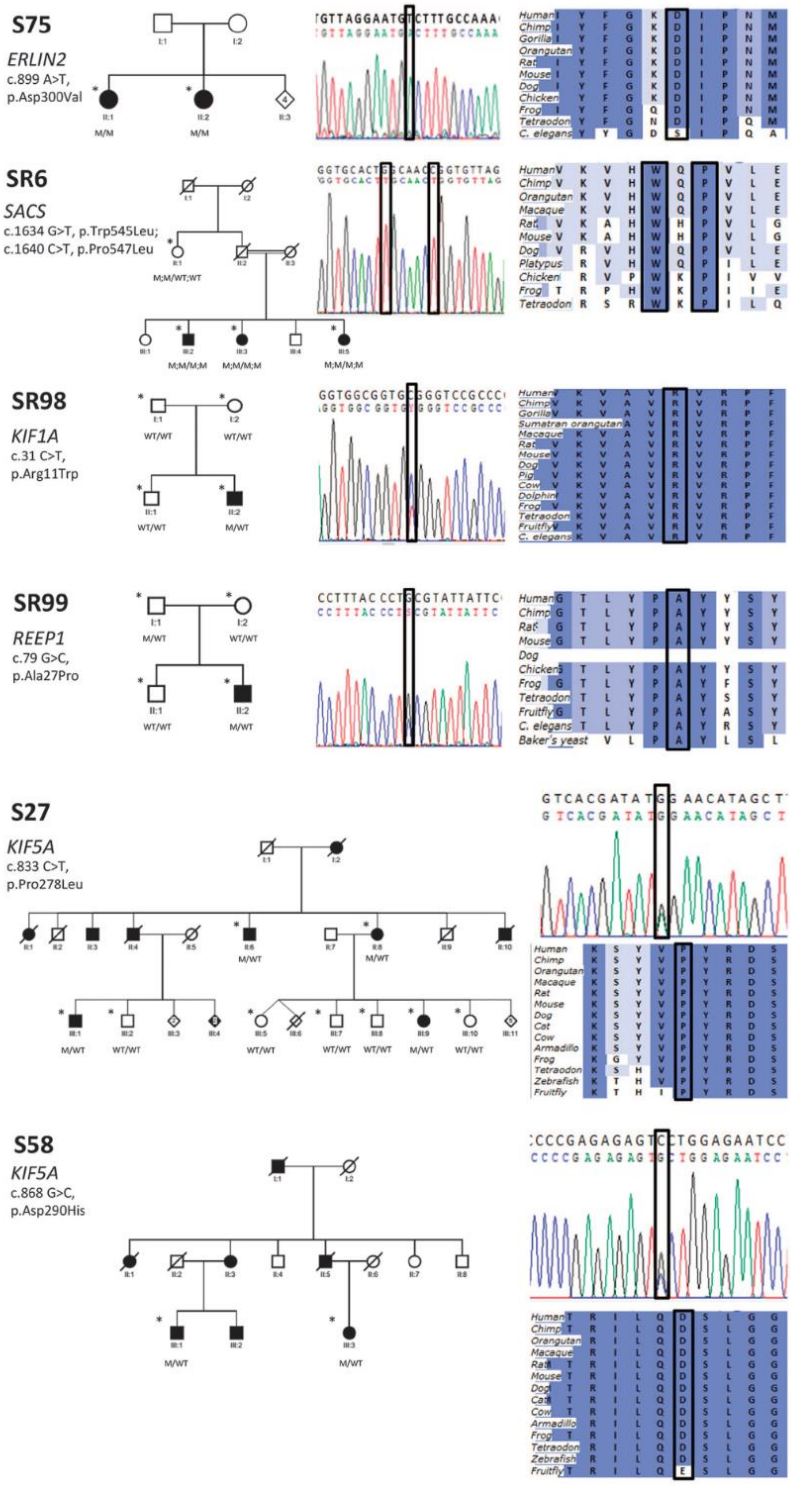


Figure 2 Pedigree with segregation analysis, electropherogram and conservation of the seven novel missense variants. A black circles and squares indicate affected individuals with HSP; an asterisk in a symbol indicates family members with available DNA. Electropherogram with the position of the missense variant boxed. Conservation of the region of the altered amino acid (boxed).

Table 2 Variants of unknown significance (VUS) found in 37 families that remained without a molecular diagnosis

Family code	Chr	Genomic coordinate (GRCh37)	Gene	Accession.Version numbers	Exon[a]	Coding change	Predicted change	State	MAF EA	Nº software[b]	ClinVar acession number	Reason for exclusion
SR1	9	135202274	*SETX*	NM_015046.5	E10	c.4711G>A	p.(Val1571Met)	Het		1/5	SCV000574465	Heterozygous variant in a recessive gene
SR5	7	99702495	*AP4M1*	NG_016312.1 (NM_004722.3)	E8	c.607-2A>G	r.?; p.?	Het		5/5	SCV000574466	Heterozygous variant of a recessive gene
SR12	1	110169006	*AMPD2*	NM_001257360.1	E6	c.650G>C	p.(Arg217Pro)	Het		2/5	SCV000574467	Heterozygous variant of a recessive gene
SR20	5	10256222	*CCT5*	NM_012073.4	E4	c.487G>A	p.(Glu163Lys)	Het		2/4	SCV000574468	Heterozygous variant of a recessive gene
SR23	13	23912386	*SACS*	NM_014363.5	E10	c.5629C>T	p.(Arg1877*)	Het		nonsense	SCV000574469	Heterozygous variant of a recessive gene
	15	44890903	*SPG11*	NM_025137.3	E22	c.3818A>G	p.(Lys1273Arg)	Het	0.110	1/5	SCV000574470	Heterozygous variant of a recessive gene
SR26	3	39108050	*WDR48*	NM_001346225.1	E4	c.280T>G	p.(Ser94Ala)	Hom	0.500	1/5	SCV000574471	Segregation could not be tested in an inconclusive variant
SR29	2	198359436	*HSPD1*	NM_002156.4	E6	c.551A>G	p.(Asn184Ser)	Het	0.081	2/5	SCV000574472	Dominant with MAF
SR49	11	62459902	*BSCL2*	NM_001122955.3	E6	c.809G>T	p.(Arg270Leu)	Het		5/5	SCV000574473	Dominant, incomplete penetrance?
SR54	1	110169006	*AMPD2*	NM_001257360.1	E6	c.650G>C	p.(Arg217Pro)	Het		2/5	SCV000574467	Heterozygous variant of a recessive gene
SR60	2	202591249	*ALS2*	NM_020919.3	E19	c.3206G>A	p.(Gly1069Glu)	Het	0.073	5/5	SCV000574474	Heterozygous variant of a recessive gene
	11	65546763	*AP5B1*	NM_138368.4	E2	c.1201C>T	p.(Leu401Phe)	Het	0.499	3/4	SCV000574475	Candidate gene
SR65	11	65545875	*AP5B1*	NM_138368.4	E2	c.2089T>C	p.(Phe697Leu)	Het		2/4	SCV000574476	Candidate gene
	7	4825294	*AP5Z1*	NM_014855.2	E9	c.1111G>A	p.(Ala371Thr)	Het	0.012	3/5	SCV000574477	Heterozygous variant of a recessive gene
SR66	7	4824586	AP5Z1	NM_014855.2	E7	c.838A>C	p.(Thr280Pro)	Het		2/5	SCV000574478	Heterozygous variant of a recessive gene
SR71	5	149677373	*ARSI*	NM_001012301.3	E2	c.1114G>A	p.(Asp372Asn)	Het	0.140	1/5	SCV000574479	Heterozygous variant of a recessive gene
	11	62458862	*BSCL2*	NM_001122955.3	E7	c.895G>A	p.(Ala299Thr)	Het	0.012	5/5	SCV000574480	Dominant with MAF
SR76	2	171709291	*GAD1*	NM_000817.2	E13	c.1252G>T	p.(Val418Phe)	Het		3/5	SCV000574481	Heterozygous variant of a recessive gene
SR81	9	95491490	*BICD2*	NM_001003800.1	E2	c.269A>G	p.(Lys90Arg)	Het	0.581	1/5	SCV000574482	Heterozygous variant of a recessive gene
SR92	9	95491490	*BICD2*	NM_001003800.1	E2	c.269A>G	p.(Lys90Arg)	Het	0.581	1/5	SCV000574482	Heterozygous variant of a recessive gene
	16	89619540	*SPG7*	NM_003119.2	E14	c.1933T>A	p.(Ser645Thr)	Het	0.070	2/5	SCV000574483	Heterozygous variant of a Recessive Gene
S19	17	4925675	*KIF1C*	NM_006612.5	E22	c.2299G>A	p.(Gly767Arg)	Het	0.245	4/5	SCV000574484	Heterozygous variant of a recessive gene
S20	11	62458862	*BSCL2*	NM_001122955.3	E7	c.895G>A	p.(Ala299Thr)	Het	0.012	5/5	SCV000574480	Dominant with MAF
	17	4908241	*KIF1C*	NM_006612.5	E13	c.1111G>A	p.(Ala371Thr)	Het	0.430	4/5	SCV000574485	Heterozygous variant of a recessive gene
S21	17	4925967	*KIF1C*	NM_006612.5	E22	c.2591G>A	p.(Arg864His)	Het	0.047	4/5	SCV000574486	Heterozygous variant of a recessive gene
	19	35793497	*MAG*	NM_002361.3	E7	c.1117A>C	p.(Ser373Arg)	Het	0.012	4/5	SCV000574487	Heterozygous variant of a recessive gene
S22	13	23929378	*SACS*	NM_014363.5	E8	c.1373C>T	p.(Thr458Ile)	Het	0.337	2/4	SCV000574488	Heterozygous variant of a Recessive Gene
S24	1	114442497	*AP4B1*	NG_031901.1 (NM_006594.3)	E6	c.1114 +29G>C	r.?; p.?	Het		4/5	SCV000574489	Heterozygous variant of a recessive gene
S32	11	63885330	*FLRT1*	NM_013280.4	E2	c.1591G>A	p.(Asp531Asn)	Het		5/5	SCV000574490	Heterozygous variant of a Recessive Gene
S35	19	35802850	*MAG*	NM_002361.3	E10	c.1646C>T	p.(Ser549Leu)	Het		2/5	SCV000574491	Heterozygous variant of a recessive gene
S37	2	202582890	*ALS2*	NM_020919.3	E24	c.3746T>C	p.(Phe1249Ser)	Het		4/5	SCV000574492	Heterozygous variant of a recessive gene
	12	57881856	*MARS*	NG_034077.1 (NM_004990.3)	E1	c.-18C>T	r.?; p.?	Het	0.020	3/5	SCV000574493	Heterozygous variant of a recessive gene
	15	23086364^ 23086365	*NIPA1*	NM_144599.4	E2	c.42_47dup	p.(Ala15_Ala16dup)	Het		Inframe insertion	SCV000574494	Segregation could not be tested in an inconclusive variant
S44	11	65546173	*AP5B1*	NM_138368.4	E2	c.1791G>C	p.(Gln597His)	Het	0.012	4/4	SCV000574495	Candidate gene
	9	95482883	*BICD2*	NM_001003800.1	E4	c.761A>G	p.(Lys254Arg)	Het		3/5	SCV000574496	Heterozygous variant of a recessive gene
	X	73641849	*SLC16A2*	NM_006517.3	E1	c.599C>A	p.(Ser200Tyr)	Het		1/4	SCV000574497	Heterozygous variant of a recessive gene
S47	13	23904298	*SACS*	NM_014363.5	E10	c.13717A>C	p.(Asn4573His)	Het	0.465	3/4	SCV000574498	Heterozygous variant of a recessive gene

Table 2 (Continued)

Family code	Chr	Genomic coordinate (GRCh37)	Gene	Accession.Version numbers	Exon[a]	Coding change	Predicted change	State	MAF EA	Nº software[b]	ClinVar accession number	Reason for exclusion
S51	11	65546763	*AP5B1*	NM_138868.4	E2	c.1201C>T	p.(Leu401Phe)	Het	0.499	3/4	SCV000574475	Candidate gene
	17	4926868	*KIF1C*	NM_006612.5	E23	c.2734C>T	p.(Arg912Trp)	Het	0.128	3/5	SCV000574499	Heterozygous variant of a recessive gene
S52	19	35790493	*MAG*	NM_002361.3	E5	c.452C>T	p.(Ala151Val)	Het	0.430	3/5	SCV000574500	Heterozygous variant of a recessive gene
S55	9	135203756	*SETX*	NM_015046.5	E10	c.3229G>A	p.(Asp1077Asn)	Het	0.093	3/5	SCV000574501	Heterozygous variant of a recessive gene
S56	7	4830418	*AP5Z1*	NM_014855.2	E16	c.2053C>T	p.(Arg685Cys)	Het	0.012	4/5	SCV000574502	Heterozygous variant of a recessive gene
S62	19	35790493	*MAG*	NM_002361.3	E5	c.452C>T	p.(Ala151Val)	Het	0.430	3/5	SCV000574500	Heterozygous variant of a recessive gene
S64	5	149677031	*ARSI*	NM_001012301.3	E2	c.1456C>T	p.(Arg486Cys)	Het	0.012	4/5	SCV000574503	Heterozygous variant of a recessive gene
	14	68264796	*ZFYVE26*	NM_015346.3	E11	c.2183G>A	p.(Arg728Gln)	Het		1/5	SCV000574504	Heterozygous variant of a recessive gene
S73	1	110170804	*AMPD2*	NM_001257360.1	E11	c.1342G>A	p.(Val448Met)	Het	0.012	2/5	SCV000574505	Heterozygous variant of a recessive gene
	9	135203159	*SETX*	NM_015046.5	E10	c.3826C>G	p.(Gln1276Glu)	Het	0.023	1/5	SCV000574506	Heterozygous variant of a recessive gene
S74	7	4830379	*AP5Z1*	NM_014855.2	E16	c.2014G>A	p.(Glu672Lys)	Het	0.012	4/5	SCV000574507	Heterozygous variant of a recessive gene
S76	16	89611173	*SPG7*	NM_003119.2	E10	c.1442C>T	p.(Thr481Met)	Het		5/5	SCV000574508	Heterozygous variant of a recessive gene
S86	7	4821291	*AP5Z1*	NM_014855.2	E3	c.272G>A	p.(Arg91Gln)	Het		3/5	SCV000574509	Heterozygous variant of a recessive gene
	15	44914091	*SPG11*	NM_025137.3	E14	c.2486T>G	p.(Val829Gly)	Het		3/5	SCV000574510	Heterozygous variant of a recessive gene
S88	12	57960979	*KIF5A*	NM_004984.2	E7	c.572G>A	p.(Arg191His)	Hom		5/5	SCV000574511	Homozygosity in a dominant gene

[a]Exon Number according to Alamut software v2.9.0 that used the NCBI RefSeq numbering scheme (along with simple systematic numbering) on the GRCh37 (hg19) assembly.
[b]Number of Software predicting a functional effect of all the prediction software for the missense and possible splice-causing variants; Chr, Chromosome; MAF EA, Minor Allele Frequency in the European American population from Exome Variant Server (http://evs.gs.washington.edu/EVS/); Het, Heterozygous; Hom, Homozygous.

agreement with Goizet and colleagues.[28] We also identified three families with *SPG7* variants (Supplementary Figure S1) with late-onset HSP and a complex phenotype: families SR75 and S26 also presenting ataxia and the sporadic case CI2 that has minor cerebellar atrophy at brain magnetic resonance imaging which confirms that the involvement of the cerebellum is part of the SPG7 core phenotype. In family SR84 (Supplementary Figure S1), presenting two known variants in the *GBA2* gene, one affecting the catalytic site and one destabilizing the protein by disturbance of the charge balance,[33] patients had an onset in the expected range (infancy or childhood) and a complex phenotype with cerebellar signs and ataxic gait (Table 3).

Interestingly, we confirmed the higher occurrence of cases with heterozygous *vs* homozygous/compound heterozygous (SPG30) variants in *KIF1A*. One patient in the SR98 family (Figure 2) carried a novel heterozygous *de novo KIF1A* variant located in the motor domain and presented a complex phenotype with mental retardation and neuropathy and an early age at onset, in the first year of life.

On the other hand, the clinical features of some of our patients extended the phenotypic spectrum of some clinico-genetic entities. This was the case of the family with the *SACS* variants (SR6, Figure 2 and Table 3), in which patients presented with spasticity and peripheral neuropathy but without ataxia, as in another single case recently reported in an Italian family.[15] In the family with a *C19orf12* variant (family SR88, Supplementary Figure S1), the phenotype was similar to SPG43 but also presenting cerebellar ataxia, mental retardation and dementia. Interestingly, the same variant was described in patients with neurodegeneration with brain iron accumulation type 4 (NBIA4),[20] extending the spectrum of phenotypes associated with this variant. Brain imaging data were not available for this patient, so we could not check for the presence of brain iron accumulation. Lastly, patients of the S75 family (Figure 2) showing *ERLIN2* variants had an early onset, at 8 and 19 years old, however with a pure phenotype, contrary to the complex phenotype previously described.[34]

DISCUSSION

Genetic diagnostic yield

The identification of the genetic cause for HSP in only 20.4% of this cohort shows that there is still a large set of genes responsible for spastic paraplegia to be uncovered or that novel inheritance modes, as in the case of *KIF1A*,[35] should be taken into account. Although most of these families had been previously screened for *SPG3*, *SPG4* and *SPG31* in the families with dominant forms[36] and for *SPG11*, *SPG15* and *SPG5* in the families with recessive transmission (unpublished data), we expected a higher frequency of diagnosed families since we were testing almost all the known genetic causes of HSP, including all the genes identified by Novarino and colleagues.[37] This frequency is very similar to that found in a previous study covering far fewer genes in patients where only *SPAST* (SPG4) variants were excluded.[38] On the other hand, our results are also very similar and comparable to a study where exome sequencing identified disease-causing variants in eight of 48 HSP families (16.6%) and potentially disease-causing variants in another eight families where previous screening for the most likely genes had been performed.[39] Taking into account the full cohort, we have found the responsible gene in 42% of the Portuguese families (81/193), or in 51% if we only take into account the families with available DNA (81/159), a frequency similar to a Greek study where only 16 HSP genes were screened.[40] The most frequently mutated genes are (Figure 4): *SPAST* (SPG4) with a frequency of 18% (28/159) and *SPG11* with a frequency of 13% (20/159). Interestingly, we found a high frequency of *KIF5A* variants, present in four out of 44 families

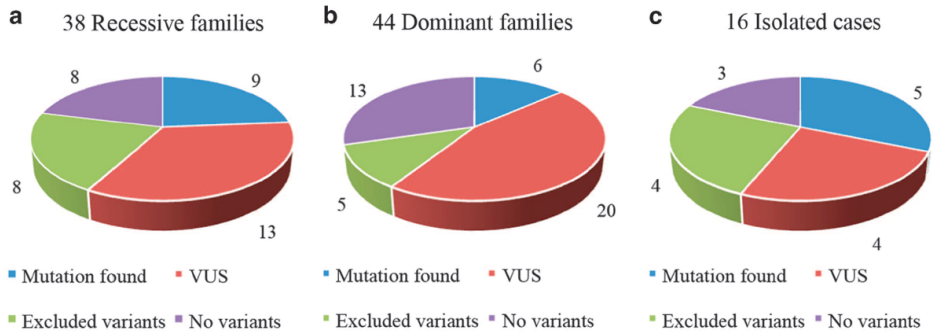


Figure 3 Distribution of the number of families in each class of variants identified. (**a**) Recessive families, (**b**) Dominant families, (**c**) Isolated cases.

with dominant inheritance (9.1%) and in 4 of the 159 families (2.5%) becoming the third most frequently mutated gene in dominant forms. This high occurrence was also found in the Greek population where it was the second cause of disease in families with dominant transmission[40] and in the study by Warrenburg *et al* where it was found as the most frequently mutated gene (after exclusion of the most likely genes).[39] Most of the other genes tested in our cohort have accounted for single families until now, as is the case of *REEP2* (SPG72)[41] and in the case of a family that was negative in this screening but after exome analysis proved to segregate *ALDH18A1* (SPG9) variants.[7] In line with our results showing that most of the new HSP genes are rarely involved, the yield of positive results obtained with the two panels, the second containing 30 additional HSP genes, was very similar. We cannot exclude that a variant in the 30 additional genes added to the second panel may explain some of the 17 cases without diagnosis (of the 25) tested with the first panel. This is however very unlikely given (1) the low relative frequencies of these genes,[5,37] (2) their involvement in recessive forms while 10 of 17 of our patients were associated with dominant inheritance of the disease, (3) and because seven of these cases processed under exome sequencing more recently do not carry mutations in the genes present in the second panel (unpublished data).

Our relatively high percentage of diagnosis among the isolated cases (5 out of 16) is also an interesting finding that shows how important it is to test these patients that frequently are not considered in genetic analyses due to the absence of additional family members to confirm diagnosis.

This study also demonstrates that a panel strategy is also capable of detecting large rearrangements, as proved with the identification of a large deletion, although at the homozygous state, encompassing exons 12–14 in *SPG11* in family SR17. This was possible by the analysis of the coverage, where we compared the mean coverage of all patients tested with each other taking into account the average coverage of each one.

Towards an unbiased approach

Our results also highlight the importance of testing all families for the same set of genes, regardless the presumed inheritance mode. This is important as sometimes the transmission mode is unclear in the pedigree because of its size or censured consanguinity, and because there is an increasing number of HSP genes associated with different modes of transmission (eg, *BICD2*, *KIF1C* and *KIF1A*), sometimes associated with a different disease presentation (eg, *SPG7*). We have some examples in our cohort of families with an apparent recessive mode of transmission (including sporadic cases) in whom we found a variant in a gene associated with a dominant form. This can happen when the disease is caused by a '*de novo*' variant, as in families SR4 (Supplementary Figure S1) and SR98 (Figure 2), and also in late-onset families in which the disease could have been missed in the older generations. The reverse is also true since we report one case with a homozygous variant in *SPG7*, usually associated with a recessive form, found in a family classified as dominant (S26, Supplementary Figure S1). This could likely be explained by censured consanguinity or by the presence of older individuals reported as affected but with a different neurological condition. Family SR98, with a variant in the motor domain of *KIF1A*, is also an example of a gene that can have recessive or dominant inheritance depending on the localization of the variant. It was recently shown by Lee and colleagues that *de novo* heterozygous variants in the motor domain of *KIF1A* can be disease-causing[35] and that this type of variants are more common than the homozygous ones, which has been later confirmed in several studies[39,42,43] including our own. In family SR99 (Figure 2), we found a new variant (p.Ala27Pro) in *REEP1* that, despite being present in the unaffected father, is likely the disease-causing variant because incomplete penetrance was already described for variants in this gene and because the same amino acid was affected in another HSP case.[44]

Need for international collaborative efforts

Our results also highlight the difficulty of reaching a conclusion as to the causative nature of the variants in 37% of the cases. Large screening studies, like ours, result in a large quantity of data that has to be interpreted, a task that is not always easy even with all the available software. This is especially difficult in HSP since the disease can have different inheritance modes. An example of the difficulty in interpretation without functional data is illustrated by family SR6 (Figure 2), where we found two new variants, within six bases of distance, that segregate together, and we were not able to demonstrate which one was deleterious, although the predictions on their effect on protein function is stronger in one of them.

It is highly likely that in our 37 families with VUS, some of these VUS could be variants with new inheritance forms for already known genes. This could be the case in family SR88, where we found a *KIF5A* variant present at the homozygous state in patients in two generations of a family with no reported consanguinity (Supplementary Figure S2). This might represent the first autosomal recessive case due to *KIF5A* variants, but this cannot be ascertained in the absence of functional evidence or additional cases segregating in autosomal recessive pedigrees. We also found some variants in candidate genes selected by function (interactors) for which we were not able to conclude as to

Table 3 Clinical characteristics of the 20 families with an established molecular diagnosis

Family code	Individual No. (gender)	Gene	Disease-causing variant[a]	Inheritance	Phenotype	Age at onset (years)	Age at examination in years (date)	Disease duration (years)	Spasticity LL (0–4)	Spasticity UL (0–4)	Weakness LL (0–5)	Weakness UL (0–5)	Babinski sign	Sphincter disturbances	Dysarthria	Cognitive impairment	Other signs	MRI
SR88	IV:1 (M)	*C19orf12*	c.157G>A; c.157G>A	Recessive	Complex	8	17 (2004)	9	3	1	2	4	Yes	NA	Yes and dysphagia	Yes (1) and dementia	Motor neuropathy, cerebellar ataxia, bilateral optic atrophy, proximal spastic tetraparesis, deterioration, distal atrophy, pes cavus	NA
SR97	II:2 (F)	*CYP7B1*	c.914dup; c.825T>A	Recessive	Pure	10	36 (2006)	26	NA	0	4	5	Yes	NA	No	No	Pes cavus, gait ataxia	NA
	II:3 (M)					10	26 (2006)	16	NA	0	4	5	Yes	NA	No	No	Pes cavus, gait ataxia	NA
S93	IV:7 (M)	*CYP7B1*	c.1456C>T; c.1456C>T	Recessive	Pure	40	73 (2007)	33	4	0	NA	NA	Yes	Mild	NA	NA	No	NA
	IV:13 (F)					55	67 (2007)	12	NA	0	NA	NA	Yes	NA	NA	NA		NA
SR84	II:2 (F)	*GBA2*	c.2618G>A; c.1888C>T	Recessive	Complex	7	38 (2007)	46	4	0	4	0	Yes	Slight	No	No	Dysphagia, cerebellar signs, ataxic gait, postural tremor, distal LL hypoesthesia	NA
	II:4 (M)					12	23 (2007)	31	4	0	4	0	Yes	No	No	No	NA	NA
SR75	IV:2 (M)	*SPG7*	c.1454_1462del; c.1454_1462del	Recessive	Complex	22	34 (1999)	12	NA	NA	Yes	NA	NA	NA	Yes	No	Progressive ataxia at onset	Cerebellar atrophy (cortical and diffuse)
SR16	III:4 (F)	*SPG11*	c.733_734del; c.733_734del	Recessive	Complex with neuropathy	18	37 (1993)	19	4	3	2	2	NA	Yes	Yes (pseudobulbar)	Yes	NA	NA
	III:6 (M)					18	32 (1996)	14	4	2	2	3	Yes	No	Yes (Pseudobulbar and dysphagia)	Yes	NA	NA
	III:7 (M)					18	NA	NA	NA	NA	NA	NA	NA	No	NA	No	NA	NA
SR82	V:1 (M)	*ZFYVE26*	c.6296dup; c.6296dup	Recessive	Complex	14	25 (2004)	11	4	NA	4	5	Yes	Yes	Yes	Yes (1)		Abnormal TCC, bilateral cortical atrophy
S31	IV:8 (F)	*KIF5A*	c.610C>T	Dominant	Pure	14	51 (2006)	37	NA	0	4	5	Yes	NA	No	No		NA
S41	IV:2 (F)	*KIF5A*	c.839G>A	Dominant	Pure	20	64 (2008)	44	4	0	4	NA	Yes	NA	NA	NA		NA
	V:5 (F)					20	34 (2008)	12	1	0	NA	NA	Yes	NA	NA	NA		NA
	VI:1 (F)					10	19 (2008)	9	0	0	NA	NA	Yes	NA	NA	NA		NA
S61	II:5 (F)	*NIPA1*	c.316G>A	Dominant	Pure	9	58 (2005)	49	4	0	NA	NA	Yes	NA	NA	NA		NA
	III:2 (M)					5	32 (2005)	27	3	0	NA	NA	Yes	NA	NA	NA		NA
	IV:1 (F)					2	12 (2005)	10	NA	NA	NA	NA	Yes	NA	NA	NA		NA
S26	IV:3 (M)	*SPG7*	c.1447C>T; c.1447C>T	classified as dominant	Complex	40	55 (1998)	15	2	0	NA	NA	Yes	NA	NA	NA		Slight cerebellar atrophy
	IV:5 (M)					40	60 (2008)	20	1	0	4	5	Yes	NA	NA	NA	Ataxia	
SR98	IV:2 (M)	*KIF1A*	c.31C>T	Isolated case	Complex	1	13 (2007)	12	NA	0	5	5	No	NA	No	Yes (2)	Neuropathy	Normal
SR4	IV:3 (M)	*SPAST*	c.1133T>A	Isolated case	Pure	23	48 (1995)	25	NA	NA	NA	NA	Yes	No	NA	No	Pes cavus	NA
CI3	II:1 (M)	*SPG7*	c.1529C>T; c.376G>C	Isolated case	Complex	57	60 (2006)	3	1	0	5	5	No	No	No	No		Slight cerebellar atrophy

Table 3 (Continued)

Family code	Individual No. (gender)	Gene	Disease-causing variant[a]	Inheritance	Phenotype	Age at onset (years)	Age at examination in years (date)	Disease duration (years)	Spasticity LL (0–4)	Spasticity UL (0–4)	Weakness LL (0–5)	Weakness UL (0–5)	Babinski sign	Sphincter disturbances	Dysarthria	Cognitive impairment	Other signs	MRI
SR17	III:3 (M)	*SPG11*	c.(2244+1_2245-1)_(2620+1_2621-1)del; c.(2244+1_2245-1)_(2620+1_2621-1)del	Isolated Case	Complex	13	25 (1993)	12	3	0	4	5	Yes	No	No	Yes (1)	Pes cavus	NA
S75	II:1 (F)	*ERLIN2*	c.899A>T; c.899A>T	Recessive	Pure	8	31 (2001)	NA	NA	NA	NA	NA	NA	NA	NA	NA		
	II:2 (F)					19	24 (2001)	19	NA	NA	NA	NA	NA	NA	NA	NA	No	
SR6	IV:6 (M)	*SACS*	c.[1640C>T, 1634G>T]; c.[1640C>T, 1634G>T]	Recessive	Complex	6	55 (1996)	47	3	0	4	0	Yes	No	No	No	Pes cavus	NA
	IV:9 (F)					22	49 (1996)	27	2	0	4	0	Yes	NA	No	No	Pes cavus, neuropathy	NA
	IV:13 (F)					31	44 (1996)	13	2	0	4	0	Yes	NA	Yes	No	Pes cavus	Encephalic atrophy
S27	III:7 (M)	*KIF5A*	c.833C>T	Dominant	Complex	21	74 (2008)	53	1	0	4	5	NA	NA	No	No mental retardation, dementia probably of other etiology	Pes cavus	NA
	III:9 (F)					43	71 (2008)	28	0	0	2	5	NA	NA	Yes	No	Ataxia and extrapyramidal signs, parkinsonian syndrome only akinetic without spasticity	NA
	IV:3 (F)						47 (2008)		2	0	5	5	No	No	No	No		NA
S58	III:2 (M)	*KIF5A*	c.868G>C	Dominant	Pure	25	54 (2003)	29	NA	NA	NA	NA	NA	NA	NA	No	Hammer toes	NA
	III:6 (F)					18	44 (2004)	26	NA	0	4	5	Yes	No	No	No		NA
SR99	IV:2 (M)	*REEP1*	c.79G>C	Isolated case	Pure	2	16 (2010)	14	2	0	5	5	Yes	No	No	No		Normal

Abbreviations: M, male; F, female; NA, not available; MRI, magnetic resonance imaging.
For spasticity a scale of 0–4 according to the Modified Ashworth Scale was used. Weakness was graded from 0 to 5 using the Medical Research Council Scale for Muscle Strength.
[a]Accession.Version numbers available on Table 1.

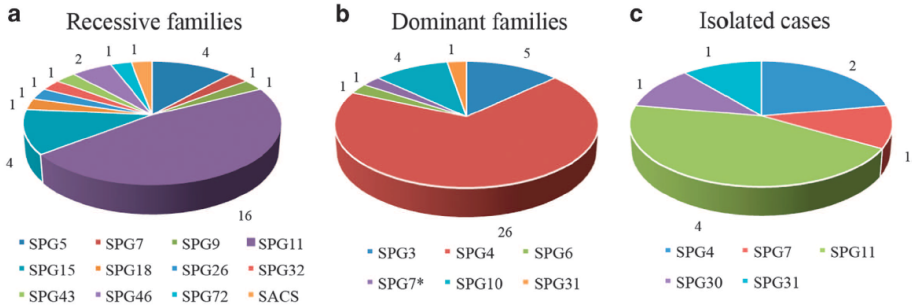


Figure 4 Distribution of the disease-causing variants found in 81 families from the full cohort of Portuguese cases.# (**a**) Recessive families, (**b**) Dominant families, (**c**) Isolated cases. *Case described in this paper of a family classified with dominant inheritance but with a homozygous *SPG7* variant. # Include mutations found in previously screened genes.[7,36,41]

their causative effect due to the insufficient number of family members to validate their co-transmission (families SR65 and S86, Supplementary Figure S2), and the absence of cell lines to demonstrate the disruption of the AP5 complex by biochemical experiments.[45] We observed a variant in a gene associated with dominant transmission, *BSCL2*, that is in a conserved region, is predicted to affect function by five software and is in the same domain as the two published variants (with reports of incomplete penetrance) but, since it is present in the non-affected mother and there are no other known carriers, we were not able to conclude as to the causative effect of this variant (case SR49, Supplementary Figure S2). Absence of family members to confirm segregation of a homozygous variant in the *WDR48* was also an issue in family SR26 (Supplementary Figure S2). This family presents a complex phenotype with motor neuropathy that is in accordance with the reported clinical features,[37] although with a later age at onset (20 and 22 years old *vs* 1-year-old), and only functional studies will allow to prove its causative role. Therefore, the availability of relatives to test segregation still remains crucial to conclude on a causative variant, and diagnosis in HSP should not rely only on the clinical/genealogical presentation because the clinical phenotype and full mutational spectrum associated with each gene are continuously expanding.[5]

In conclusion, this comprehensive study of a large cohort of 98 families of homogeneous origin (only Portuguese families) allowed us to conclude on the frequency of all the less common HSP genes. Also, it highlights the fact that although a high number of genes have been already identified in this condition, a large percentage of cases remain without molecular diagnosis, some of them due to the absence of extensive screening. Gene panel strategy is a cost-effective way to screen HSP families and probably a small panel like ours with 34 genes could give a diagnosis to almost half of the families if used as a first approach. The use of an extensive panel of genes like our second panel, including the more recent and less frequently mutated genes such as VCP[39,46] would only slightly increase the number of diagnosis. This suggests that performing exome sequencing in the families without diagnosis in the most frequent genes (34) is probably a good strategy in HSP.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We would like to thank the participating patients and families and also the clinicians as well as Federica Barrecca, Emeline Mundwiller and Delphine Bouteiller for their help in the screening. This study was funded in part by the FCT (Fundação para a Ciência e Tecnologia, grant FCT-ANR/BEX-GMG/0008/2013, to IA) and by the French National Agency for Research (SPATAX-QUEST, to GS), the European Union (Omics call of the 7th framework program, NEUROMICS, to AB; the E-rare program, NEUROLIPID and PREPARE, to GS), the VERUM Foundation (to AB and GS) and benefited from the program 'Investissements d'Avenir' ANR-10-IAIHU-06 (to the ICM). SM is the recipient of a fellowship (SFRH/BD/87189/2012) from the FCT supported by POPH/FSE funding.

1 Fink JK: Hereditary spastic paraplegia: clinico-pathologic features and emerging molecular mechanisms. *Acta Neuropathol* 2013; **126**: 307–328.
2 Ruano L, Melo C, Silva MC, Coutinho P: The global epidemiology of hereditary ataxia and spastic paraplegia: a systematic review of prevalence studies. *Neuroepidemiology* 2014; **42**: 174–183.
3 Coutinho P, Ruano L, Loureiro JL *et al*: Hereditary ataxia and spastic paraplegia in Portugal: a population-based prevalence study. *JAMA Neurol* 2013; **70**: 746–755.
4 Lo Giudice T, Lombardi F, Santorelli FM, Kawarai T, Orlacchio A: Hereditary spastic paraplegia: clinical-genetic characteristics and evolving molecular mechanisms. *Exp Neurol* 2014; **261**: 518–539.
5 Tesson C, Koht J, Stevanin G: Delving into the complexity of hereditary spastic paraplegias: how unexpected phenotypes and inheritance modes are revolutionizing their nosology. *Hum Genet* 2015; **134**: 511–538.
6 Rinaldi C, Schmidt T, Situ AJ *et al*: Mutation in CPT1C Associated With Pure Autosomal Dominant Spastic Paraplegia. *JAMA Neurol* 2015; **72**: 561–570.
7 Coutelier M, Goizet C, Durr A *et al*: Alteration of ornithine metabolism leads to dominant and recessive hereditary spastic paraplegia. *Brain* 2015; **138**: 2191–2205.
8 Gan-Or Z, Bouslam N, Birouk N *et al*: Mutations in CAPN1 cause autosomal-recessive hereditary spastic paraplegia. *Am J Hum Genet* 2016; **98**: 1038–1046.
9 Yang Y, Liu W, Fang Z *et al*: A newly identified missense mutation in FARS2 causes autosomal-recessive spastic paraplegia. *Hum Mutat* 2016; **37**: 165–169.
10 Estrada-Cuzcano A, Martin S, Chamova T *et al*: Loss-of-function mutations in the ATP13A2/PARK9 gene cause complicated hereditary spastic paraplegia (SPG78). *Brain* 2017; **140**: 287–305.
11 Josifova DJ, Monroe GR, Tessadori F *et al*: Heterozygous KIDINS220/ARMS nonsense variants cause spastic paraplegia, intellectual disability, nystagmus, and obesity. *Hum Mol Genet* 2016; **25**: 2158–2167.
12 Ozes B, Karagoz N, Schule R *et al*: Pla2g6 mutations associated with a continuous clinical spectrum from neuroaxonal dystrophy to hereditary spastic paraplegia. *Clin Genet* 2017; e-pub ahead of print 13 March 2017; doi:10.1111/cge.13008.
13 Silva MC, Coutinho P, Pinheiro CD, Neves JM, Serrano P: Hereditary ataxias and spastic paraplegias: methodological aspects of a prevalence study in Portugal. *J Clin Epidemiol* 1997; **50**: 1377–1384.
14 Wakil SM, Ramzan K, Abuthuraya R *et al*: Infantile-onset ascending hereditary spastic paraplegia with bulbar involvement due to the novel ALS2 mutation c.2761C>T. *Gene* 2014; **536**: 217–220.
15 Gregianin E, Vazza G, Scaramel E *et al*: A novel SACS mutation results in non-ataxic spastic paraplegia and peripheral neuropathy. *Eur J Neurol* 2013; **20**: 1486–1491.
16 Strong MJ, Gordon PH: Primary lateral sclerosis, hereditary spastic paraplegia and amyotrophic lateral sclerosis: discrete entities or spectrum? *Amyotroph Lateral Scler Other Motor Neuron Disord* 2005; **6**: 8–16.
17 Nisipeanu P, Kuritzky A, Korczyn AD: Familial levodopa-responsive parkinsonian-pyramidal syndrome. *Mov Disord* 1994; **9**: 673–675.

18 Furuta N, Ikeda M, Hirayanagi K, Fujita Y, Amanuma M, Okamoto K: A novel GJA1 mutation in oculodentodigital dysplasia with progressive spastic paraplegia and sensory deficits. *Intern Med* 2012; **51**: 93–98.
19 Dale RC, Gornall H, Singh-Grewal D, Alcausin M, Rice GI, Crow YJ: Familial Aicardi-Goutieres syndrome due to SAMHD1 mutations is associated with chronic arthropathy and contractures. *Am J Med Genet A* 2010; **152A**: 938–942.
20 Hartig MB, Iuso A, Haack T *et al*: Absence of an orphan mitochondrial protein, c19orf12, causes a distinct clinical subtype of neurodegeneration with brain iron accumulation. *Am J Hum Genet* 2011; **89**: 543–550.
21 Goizet C, Boukhris A, Durr A *et al*: CYP7B1 mutations in pure and complex forms of hereditary spastic paraplegia type 5. *Brain* 2009; **132**: 1589–1600.
22 Hammer MB, Eleuch-Fayache G, Schottlaender LV *et al*: Mutations in GBA2 cause autosomal-recessive cerebellar ataxia with spasticity. *Am J Hum Genet* 2013; **92**: 245–251.
23 Martin E, Schule R, Smets K *et al*: Loss of function of glucocerebrosidase GBA2 is responsible for motor neuron defects in hereditary spastic paraplegia. *Am J Hum Genet* 2013; **92**: 238–244.
24 McDermott CJ, Dayaratne RK, Tomkins J *et al*: Paraplegin gene analysis in hereditary spastic paraparesis (HSP) pedigrees in northeast England. *Neurology* 2001; **56**: 467–471.
25 Stevanin G, Santorelli FM, Azzedine H *et al*: Mutations in SPG11, encoding spatacsin, are a major cause of spastic paraplegia with thin corpus callosum. *Nat Genet* 2007; **39**: 366–372.
26 Goizet C, Boukhris A, Maltete D *et al*: SPG15 is the second most common cause of hereditary spastic paraplegia with thin corpus callosum. *Neurology* 2009; **73**: 1111–1119.
27 Tessa A, Silvestri G, de Leva MF *et al*: A novel KIF5A/SPG10 mutation in spastic paraplegia associated with axonal neuropathy. *J Neurol* 2008; **255**: 1090–1092.
28 Goizet C, Boukhris A, Mundwiller E *et al*: Complicated forms of autosomal dominant hereditary spastic paraplegia are frequent in SPG10. *Hum Mutat* 2009; **30**: E376–E385.
29 Chen S, Song C, Guo H *et al*: Distinct novel mutations affecting the same base in the NIPA1 gene cause autosomal dominant hereditary spastic paraplegia in two Chinese families. *Hum Mutat* 2005; **25**: 135–141.
30 Tang B, Zhao G, Xia K *et al*: Three novel mutations of the spastin gene in Chinese patients with hereditary spastic paraplegia. *Arch Neurol* 2004; **61**: 49–55.
31 Brugman F, Scheffer H, Wokke JH *et al*: Paraplegin mutations in sporadic adult-onset upper motor neuron syndromes. *Neurology* 2008; **71**: 1500–1505.
32 Sanchez-Ferrero E, Coto E, Beetz C *et al*: SPG7 mutational screening in spastic paraplegia patients supports a dominant effect for some mutations and a pathogenic role for p.A510V. *Clin Genet* 2013; **83**: 257–262.
33 Charoenwattanasatien R, Pengthaisong S, Breen I *et al*: Bacterial beta-glucosidase reveals the structural and functional basis of genetic defects in human glucocerebrosidase 2 (GBA2). *ACS Chem Biol* 2016; **11**: 1891–1900.
34 Alazami AM, Adly N, Al Dhalaan H, Alkuraya FS: A nullimorphic ERLIN2 mutation defines a complicated hereditary spastic paraplegia locus (SPG18). *Neurogenetics* 2011; **12**: 333–336.
35 Lee JR, Srour M, Kim D *et al*: De novo mutations in the motor domain of KIF1A cause cognitive impairment, spastic paraparesis, axonal neuropathy, and cerebellar atrophy. *Hum Mutat* 2015; **36**: 69–78.
36 Loureiro JL, Brandao E, Ruano L *et al*: Autosomal dominant spastic paraplegias: a review of 89 families resulting from a portuguese survey. *JAMA Neurol* 2013; **70**: 481–487.
37 Novarino G, Fenstermaker AG, Zaki MS *et al*: Exome sequencing links corticospinal motor neuron disease to common neurodegenerative disorders. *Science* 2014; **343**: 506–511.
38 Kumar KR, Blair NF, Vandebona H *et al*: Targeted next generation sequencing in SPAST-negative hereditary spastic paraplegia. *J Neurol* 2013; **260**: 2516–2522.
39 van de Warrenburg BP, Schouten MI, de Bot ST *et al*: Clinical exome sequencing for cerebellar ataxia and spastic paraplegia uncovers novel gene-disease associations and unanticipated rare disorders. *Eur J Hum Genet* 2016; **24**: 1460–1466.
40 Lynch DS, Koutsis G, Tucci A *et al*: Hereditary spastic paraplegia in Greece: characterisation of a previously unexplored population using next-generation sequencing. *Eur J Hum Genet* 2016; **24**: 857–863.
41 Esteves T, Durr A, Mundwiller E *et al*: Loss of association of REEP2 with membranes leads to hereditary spastic paraplegia. *Am J Hum Genet* 2014; **94**: 268–277.
42 Ylikallio E, Kim D, Isohanni P *et al*: Dominant transmission of de novo KIF1A motor domain variant underlying pure spastic paraplegia. *Eur J Hum Genet* 2015; **23**: 1427–1430.
43 Citterio A, Arnoldi A, Panzeri E *et al*: Variants in KIF1A gene in dominant and sporadic forms of hereditary spastic paraparesis. *J Neurol* 2015; **262**: 2684–2690.
44 Beetz C, Schule R, Deconinck T *et al*: REEP1 mutation spectrum and genotype/phenotype correlation in hereditary spastic paraplegia type 31. *Brain* 2008; **131**: 1078–1086.
45 Hirst J, Edgar JR, Esteves T *et al*: Loss of AP-5 results in accumulation of aberrant endolysosomes: defining a new type of lysosomal storage disease. *Hum Mol Genet* 2015; **24**: 4984–4996.
46 de Bot ST, Schelhaas HJ, Kamsteeg EJ, van de Warrenburg BP: Hereditary spastic paraplegia caused by a mutation in the VCP gene. *Brain* 2012; **135**: e223, author reply e224.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-nd/4.0/

© The Author(s) 2017

Supplementary Information accompanies this paper on European Journal of Human Genetics website (http://www.nature.com/ejhg)

Supplementary table and figures:

Table S1. Genes searched for in the two HSP panels

Panel 1: 34 genes	Panel 2: 70 genes	
ALS2	*ALS2*	*KIAA0196*
AP4B1	*AMPD2*	*KIF1A*
AP4E1	*AP4B1*	*KIF1C*
AP4M1	*AP4E1*	*KIF5A*
AP4S1	*AP4M1*	*L1CAM*
AP5Z1	*AP4S1*	*MAG*
ATL1	*AP5B1**	*MARS*
BSCL2	*AP5M1**	*MT-ATP6*
CYP2U1	*AP5S1**	*NIPA1*
CYP7B1	*AP5Z1*	*NT5C2*
DDHD1	*ARL6IP1*	*PGAP1*
ERLIN2	*ARSI*	*PLP1*
FA2H	*ATL1*	*PNPLA6*
GBA2	*B4GALNT1*	*RAB3GAP2*
GJC2	*BICD2*	*REEP1*
HSPD1	*BSCL2*	*REEP2*
KIAA0196	*C12orf65*	*RTN2*
KIF1A	*C19orf12*	*SACS*
KIF5A	*CCT5*	*SAMHD1*
NIPA1	*CPT1C*	*SETX*
PLP1	*CYP2U1*	*SLC16A2*
PNPLA6	*CYP7B1*	*SLC33A1*
REEP1	*DDHD1*	*SPAST*
RTN2	*DDHD2*	*SPG11*
SACS	*ENTPD1*	*SPG20*
SETX	*ERLIN1*	*SPG21*
SLC16A2	*ERLIN2*	*SPG7*
SLC33A1	*FA2H*	*TECPR2*
SPAST	*FBXO7*	*TFG*
SPG11	*FLRT1*	*USP8*
SPG20	*GAD1*	*VCP*
SPG21	*GBA2*	*VPS37A*
SPG7	*GJA1*	*WDR48*
ZFYVE26	*GJC2*	*ZFR*
	HSPD1	*ZFYVE26*

*Candidate genes

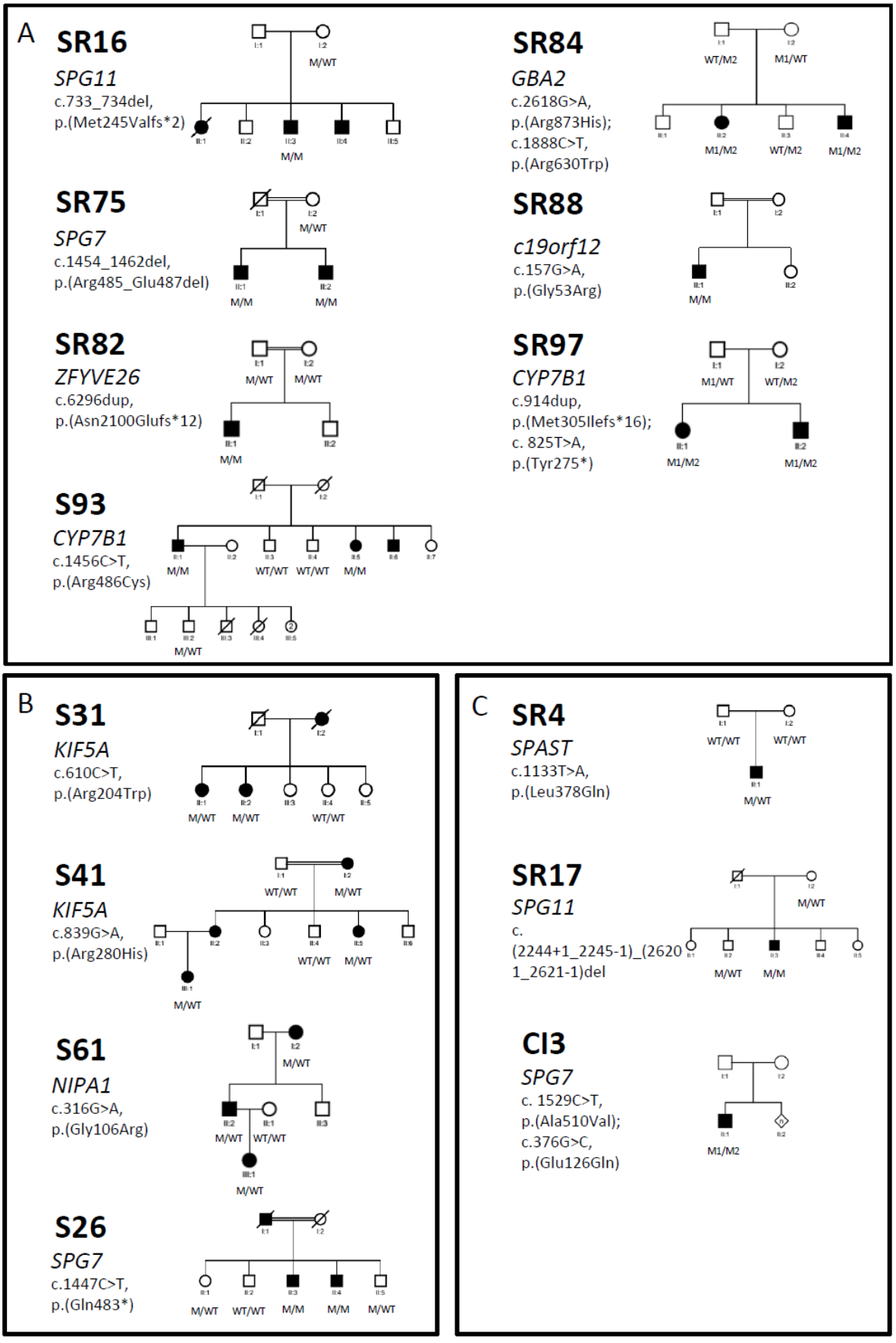
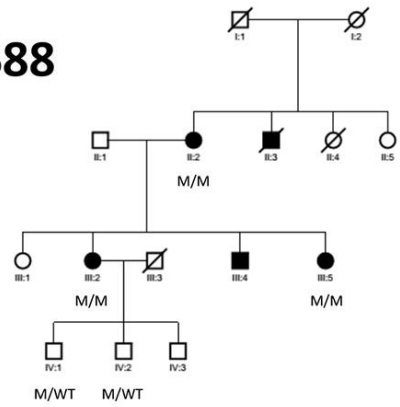
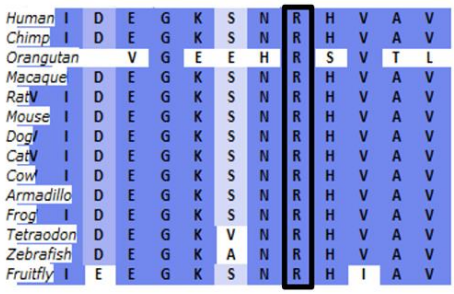


Figure S1. Segregation of the variants found in the remaining 14 families with established diagnosis. A) Families with recessive inheritance, B) Families with recessive inheritance, C) Isolated cases

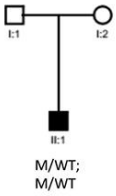
S88



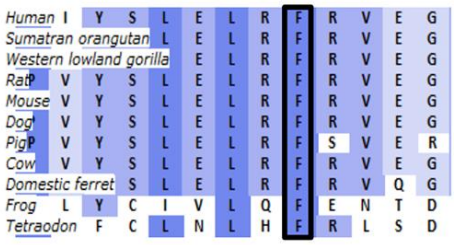
KIF5A
c.572G>A,
p.Arg191His



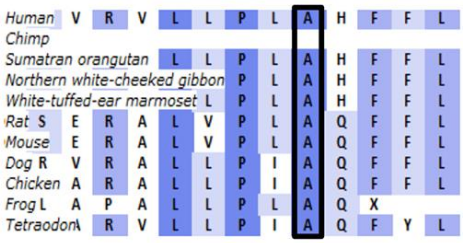
SR65



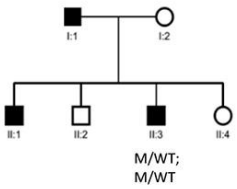
AP5B1
c.2089T>C,
p.Phe697Leu



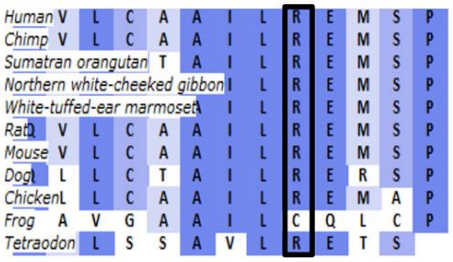
AP5Z1
c.1111G>A,
p.Ala371Thr



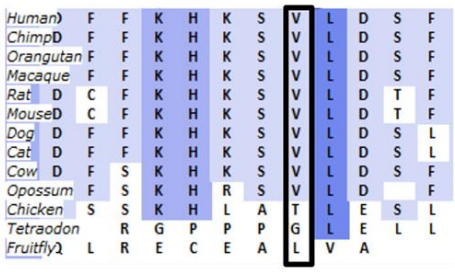
S86



AP5Z1
c.272G>A,
p.Arg91Gln



SPG11
c.2486T>G,
p.Val829Gly



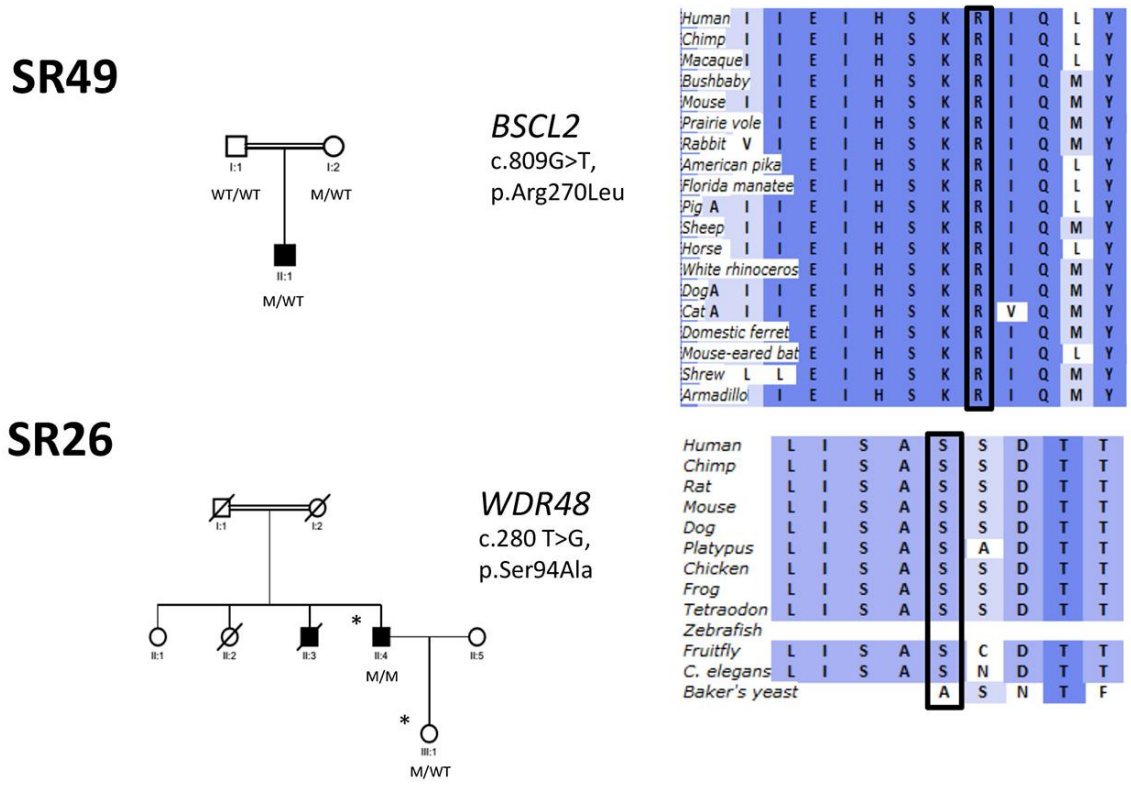


Figure S2. Segregation analysis and conservation of the variants found in the five families with a non-definitive diagnosis. A family with recessive transmission showing a variant in a gene associated with dominant form (S88), two families with possible digenism (SR65 and S86), a family with a possible dominant inheritance with incomplete penetrance (SR49) and a family with an inconclusive homozygous variant (SR26). Black circles and squares indicate affected individuals with HSP. Conservation of the region of the altered amino acid (boxed).

4.2 Search for new HSP causing genes

The first part of our work entitled “search for molecular diagnosis in already identified genes” allowed us to provide a molecular diagnosis to 20 families of our cohort. In order to be able to achieve a diagnosis in additional families, the second strategy selected was the search for new disease causing genes by exome sequencing. We selected twelve of the most informative families (of the 80 without diagnosis): seven AR and five AD families for whole exome sequencing (WES). As exome analysis is a strategy that results in the identification of a huge amount of variants, even after all the bioinformatics analysis that filters and reduces this number, sometimes there is the need to complement this analysis with additional methods that help to focus the search. Therefore, for the five families with autosomal dominant inheritance we also performed genome-wide genotyping and linkage analysis in order to reduce the chromosomal area to focus in when searching for the causing variants in the exome analysis. Three of the autosomal recessive families already had linkage data from a previous study of 300 autosomal microsatellites (Stevanin et al., 2007b).
In these twelve families, we found candidate genes in 9 families, of which 5 only had one candidate gene after bioinformatics analysis. We selected two of them for further functional studies (*AIMP1* and *AKAP6*). A third candidate gene, *ALDH18A1,* had recently been found as a disease-causing gene in other families from the SPATAX cohort, so the Portuguese family with variants in this gene (family SR45) was included in a common publication (article 2).

This work was described as part of a paper from Marie Coutelier published in the Brain Journal (article 2 presented in 4.2.1) and one brief report (submitted) describing the new HSP gene, *AIMP1* corresponding to article 3 (presented in 4.2.2). The preliminary data obtained for *AKAP6* is presented in 4.2.3. The remaining results from the exome analysis are presented in 4.2.4, which include a strong candidate gene, *SUCO*, also selected for functional study.

Article 2: Coutelier M, Goizet C, Durr A, Habarou F, Morais S, Dionne-Laporte A, Tao F, Konop J, Stoll M, Charles P, Jacoupy M, Matusiak R, Alonso I, Tallaksen C, Mairey M, Kennerson M, Gaussen M, Schule R, Janin M, Morice-Picard F, Durand C, Depienne C, Calvas P, Coutinho P, Saudubray JM, Rouleau G, Brice A, Nicholson G, Darios F, Loureiro JL, Zuchner S, Ottolenghi C, Mochel F, Stevanin G: “Alteration of ornithine metabolism leads to dominant and recessive hereditary spastic paraplegia”. Brain 2015 Aug;138(Pt 8):2191-205. doi: 10.1093/brain/awv143

Article 3: Morais S, Tesson C, Coutinho P, Brandão E, Ribeiro P, Loureiro JL, Sequeiros J, Brice A, Stevanin G, Alonso I. “AIMP1, a gene implicated in a range of phenotypes from intellectual disability to Hereditary Spastic Paraplegia”. Submitted

4.2.1 *ALDH18A1* as a causative gene for HSPs

doi:10.1093/brain/awv143 BRAIN 2015: 138; 2191–2205 | 2191

BRAIN
A JOURNAL OF NEUROLOGY

Alteration of ornithine metabolism leads to dominant and recessive hereditary spastic paraplegia

Marie Coutelier,[1,2,3,4,5,6] Cyril Goizet,[7,8] Alexandra Durr,[1,2,3,4,9] Florence Habarou,[10] Sara Morais,[1,2,3,4,6,11,12,13] Alexandre Dionne-Laporte,[14] Feifei Tao,[15] Juliette Konop,[1,2,3,4,6] Marion Stoll,[16] Perrine Charles,[9] Maxime Jacoupy,[1,2,3,4] Raphaël Matusiak,[1,2,3,4] Isabel Alonso,[11,12,13] Chantal Tallaksen,[1,2,3,4,#] Mathilde Mairey,[1,2,3,4,6] Marina Kennerson,[16] Marion Gaussen,[1,2,3,4,6] Rebecca Schule,[15,17,18] Maxime Janin,[10] Fanny Morice-Picard,[7,8] Christelle M. Durand,[7] Christel Depienne,[1,2,3,4,9] Patrick Calvas,[19] Paula Coutinho,[11,12,20] Jean-Marie Saudubray,[9] Guy Rouleau,[14,21] Alexis Brice,[1,2,3,4,9] Garth Nicholson,[16] Frédéric Darios,[1,2,3,4] José L. Loureiro,[11,20] Stephan Zuchner,[15] Chris Ottolenghi,[10] Fanny Mochel[1,2,3,4,9] and Giovanni Stevanin[1,2,3,4,6,9]

Hereditary spastic paraplegias are heterogeneous neurological disorders characterized by a pyramidal syndrome with symptoms predominantly affecting the lower limbs. Some limited pyramidal involvement also occurs in patients with an autosomal recessive neurocutaneous syndrome due to *ALDH18A1* mutations. *ALDH18A1* encodes delta-1-pyrroline-5-carboxylate synthase (P5CS), an enzyme that catalyses the first and common step of proline and ornithine biosynthesis from glutamate. Through exome sequencing and candidate gene screening, we report two families with autosomal recessive transmission of *ALDH18A1* mutations, and predominant complex hereditary spastic paraplegia with marked cognitive impairment, without any cutaneous abnormality. More interestingly, we also identified monoallelic *ALDH18A1* mutations segregating in three independent families with autosomal dominant pure or complex hereditary spastic paraplegia, as well as in two sporadic patients. Low levels of plasma ornithine, citrulline, arginine and proline in four individuals from two families suggested P5CS deficiency. Glutamine loading tests in two fibroblast cultures from two related affected subjects confirmed a metabolic block at the level of P5CS *in vivo*. Besides expanding the clinical spectrum of *ALDH18A1*-related pathology, we describe mutations segregating in an autosomal dominant pattern. The latter are associated with a potential trait biomarker; we therefore suggest including amino acid chromatography in the clinico-genetic work-up of hereditary spastic paraplegia, particularly in dominant cases, as the associated phenotype is not distinct from other causative genes.

1 INSERM, U 1127, F-75013, Paris, France
2 CNRS, UMR 7225, F-75013, Paris, France
3 Sorbonne Universités, UPMC Univ Paris 06, UMRS_1127, F-75013, Paris, France
4 Institut du Cerveau et de la Moelle épinière, ICM, F-75013, Paris, France
5 Laboratory of Human Molecular Genetics, de Duve Institute, Université catholique de Louvain, B-1200, Brussels, Belgium
6 Ecole Pratique des Hautes Etudes, F-75014, Paris, France
7 Univ. Bordeaux, Laboratoire Maladies Rares: Génétique et Métabolisme, EA4576, F-33000, Bordeaux, France
8 CHU Pellegrin, Service de Génétique Médicale, F-33000, Bordeaux, France
9 APHP, Hôpital de la Pitié-Salpêtrière, Département de Génétique, F-75013, Paris, France
10 Metabolic Biochemistry Lab, Necker-Enfants Malades Hospital, APHP, F-75015; and University Paris Descartes, F-75006, Paris, France

Received February 7, 2015. Revised March 17, 2015. Accepted April 4, 2015. Advance Access publication May 30, 2015
© The Author (2015). Published by Oxford University Press on behalf of the Guarantors of Brain. All rights reserved.
For Permissions, please email: journals.permissions@oup.com

11 UnIGENe, Instituto de Biologia Molecular e Celular (IBMC), Universidade do Porto, P-4150, Porto, Portugal
12 Instituto de Investigação e Inovação em Saúde, Universidade do Porto, P-4150, Porto, Portugal
13 Instituto de Ciências Biomédicas de Abel Salazar (ICBAS), Universidade do Porto, P-4150, Porto, Portugal
14 Montreal Neurological Institute and Hospital, McGill University, Montreal, QC H3A 2B4, Canada
15 Dr John T. Macdonald Foundation Department of Human Genetics and John P. Hussman Institute for Human Genomics, University of Miami Miller School of Medicine, Miami, FL 33136, USA
16 Northcott Neuroscience Laboratory, ANZAC Research Institute; Molecular Medicine Laboratory, Concord Hospital; Sydney Medical School University of Sydney, NSW 2138, Sydney, Australia
17 Centre for Neurology and Hertie Institute for Clinical Brain Research, Eberhard-Karls-University, G-72074, Tübingen, Germany
18 German Centre of Neurodegenerative Diseases (DZNE), Eberhard-Karls-University, G-72074, Tübingen, Germany
19 Fédération de Neurologie et Service de Génétique Médicale, CHU de Toulouse, Hôpital Purpan, F-31059, Toulouse, France
20 Serviço de Neurologia, Centro Hospitalar de Entre o Douro e Vouga, P-4520-211, Santa Maria da Feira, Portugal
21 Department of Neurology and Neurosurgery, McGill University, Montreal, QC H3A 2B4, Canada

[#]Present address: Department of Neurology, Oslo University Hospital; and Faculty of Medicine, Oslo University, Oslo, Norway

Correspondence to: Giovanni Stevanin,
Institut du Cerveau et de la Moelle épinière,
CHU Pitié-Salpêtrière,
47 Bd de l'Hôpital,
75013, Paris,
France
E-mail: giovanni.stevanin@upmc.fr

Keywords: ornithine; hereditary spastic paraplegia; delta-1-pyrroline-5-carboxylate synthase; *ALDH18A1*; citrulline

Abbreviations: HSP = hereditary spastic paraplegia; P5CS = delta-1-pyrroline-5-carboxylate synthase

Introduction

Hereditary spastic paraplegias (HSPs) are a family of neurodegenerative disorders mainly affecting the corticospinal tract. Clinically, they are characterized by a pyramidal syndrome with paraparesis and spasticity, primarily detectable in the lower limbs and in the gait, and classified as either pure or complex, according to the absence or presence of accompanying neurological or extra-neurological signs (Blackstone, 2012; Finsterer *et al.*, 2012). Genetically, they can follow all modes of inheritance: autosomal dominant, autosomal recessive, X-linked and mitochondrial. Many genes have already been described, and more are regularly identified (Novarino *et al.*, 2014); in this respect, next generation sequencing techniques are powerful tools to elucidate the genetic basis of HSPs, either for known gene screening or for new gene identification.

In the last decade, mutations in *ALDH18A1* have been implicated in an autosomal recessive neurocutaneous syndrome, characterized by severe developmental delay with marked cognitive impairment, associated with progeroid features, cutis laxa, joint hyperlaxity, short stature, cataract and frequent microcephaly (Baumgartner *et al.*, 2000, 2005; Bicknell *et al.*, 2008; Skidmore *et al.*, 2011; Martinelli *et al.*, 2012; Zampatti *et al.*, 2012; Fischer *et al.*, 2014; Gardeitchik *et al.*, 2014; Handley *et al.*, 2014; Wolthuis *et al.*, 2014). Pyramidal signs were reported in 12 of 15 patients identified so far, but were often limited to brisk deep tendon reflexes. *ALDH18A1* encodes delta-1-pyrroline-5-carboxylate synthase (P5CS), an enzyme that catalyses the first common steps of proline and ornithine biosynthesis from glutamate (Hu *et al.*, 1999). Plasma amino acid levels were abnormal in three reported families (Baumgartner *et al.*, 2000, 2005; Martinelli *et al.*, 2012; Fischer *et al.*, 2014), with low proline, ornithine (2/3), citrulline and arginine (3/3) levels associated with mild fasting hyperammonaemia. In this paper, we report seven pedigrees with *ALDH18A1* mutations segregating in a recessive or, for the first time, a dominant inheritance mode associated with abnormal plasma amino acid levels.

Materials and methods

Patient recruitment and clinical evaluation

Families FSP410, FSP429, FSP470, FSP856 and SR45 were of French ($n = 2$), Spanish ($n = 1$), Italian ($n = 1$) or Portuguese ($n = 1$) ancestry and had been identified as part of the SPATAX (https://spatax.wordpress.com/) cohort of patients with HSP. Two sporadic cases were Caucasian, from Australia/UK and USA. All patients were examined by at least one of the co-authors. Blood samples were obtained after informed and signed consent according to local ethics regulations. DNA was extracted using a standard protocol.

Whole genome mapping and exome sequencing in Family FSP410

Whole genome linkage analysis was performed using Illumina LINKAGE_12 SNP microarrays. Genotypes were determined

using Beadstudio (Illumina) and analysed with MERLIN 1.0 (Abecasis *et al.*, 2002) (Supplementary Fig. 1). Exome sequencing was performed on four affected subjects (Patients 6, 8, 25 and 29; Fig. 1). Targeted capture was done with the Agilent SureSelect Human All Exon Capture V4 XT 51MB kit. Paired-end 100 bp sequencing was performed on an Illumina HiSeq2000 system. Sequence reads were trimmed using FastX and duplicate reads were removed by Picard v1.59 (http://picard.sourceforge.net). Sequence reads were aligned to the human genome reference (hg19) using the Burroughs-Wheeler algorithm v0.6.2, followed by a local realignment around indels (insertions/deletions) using the Broad Institute Genome Analysis Tool Kit (GATK v1.4; DePristo *et al.*, 2011). Variants (SNP/INDEL) were called using GATK's Unified Genotyper module and annotated with Annovar (Wang *et al.*, 2010). They were sorted according to their localization in putatively linked or not excluded loci, expected transmission mode (heterozygous in all four individuals), frequency lower than 0.1% in public databases (CompleteGenomics, http://www.completegenomics.com/public-data/69-Genomes/; EVS, http://evs.gs.washington.edu/EVS; 1000genomes, http://www.1000genomes.org/; Exome Aggregation Consortium, http://exac.broadinstitute.org/), absence in internal controls, frequency in internal database (<5%), effect on the coding sequence (non-synonymous, splice site, frameshift or nonsense), coverage above 10×, and amino acid conservation score (GERP++ > 0, (http://mendel.stanford.edu/SidowLab/downloads/gerp/). Segregation in all affected individuals was checked through Sanger sequencing according to standard protocols.

ALDH18A1 analysis in 435 exomes of index patients with hereditary spastic paraplegia

To assess the frequency of *ALDH18A1* mutations in autosomal dominant HSP, exome sequencing data of 160 index patients with autosomal dominant HSP were examined. Because of the previous description of *ALDH18A1* mutations transmitted in an autosomal recessive mode of inheritance, the search was extended to 275 index patients with autosomal recessive or sporadic HSP. Exome data were shared and are available (GEM.app database, https://genomics.med.miami.edu/) as part of an international collaborative effort (Gonzalez *et al.*, 2013).

Panel sequencing in 95 index patients with hereditary spastic paraplegia

An amplicon-based panel including all exons of *ALDH18A1* was designed in-house with Primer3Plus (Supplementary Table 1). For 95 patients with autosomal dominant and recessive HSP selected from the SPATAX cohort, mostly with either European (32/45 with available origin) or North African (12/45) ancestry, all amplicons were amplified using the Fluidigm Access Array technology (IFC Controller AX, FC1 Cycler, 48.48 Access Arrays), and sequenced on the MiSeq Illumina sequencer as paired-end 2 × 250 bp reads. The Burrows-Wheeler algorithm v0.7.8 was applied to align sequence reads to the UCSC Genome Browser hg19 version of the human genome and variants were called via the GATK software package v3.1-1 after realignment and recalibration. Variants meeting the aforementioned criteria were confirmed using Sanger sequencing, and segregation in other affected members in the family was verified, when possible. In all *ALDH18A1*-mutated index cases, mutations in 74 previously described HSP genes were excluded with a Roche/Nimblegen capture in-house panel followed by MiSeq sequencing (unpublished data).

Biochemical analyses

Plasma amino acid profiles were obtained by ion-exchange chromatography with standard ninhydrin detection (JEOL Aminotac JLC 500/V) as part of routine diagnostic work-up for Patients FSP410-13, -29, -32, FSP429-21, and FSP856-18 on blood and urine samples. The results were compared with those obtained from the first family published with *ALDH18A1* mutations (Baumgartner *et al.*, 2005) and from the full cohort of patients from the Necker Hospital (Paris) between 1995 and 2013 ($n > 53\,000$), after exclusion of patients <15 years or with a known metabolic disorder, leading to a final set of 5023 patients. Age normalization was performed by locally weighted regression (loess) on a reference hospital population of 5043 individuals (data not shown).

Dermal fibroblasts from Patients FSP410-29 and FSP410-32 were grown in Dulbecco's modified Eagle's medium with 10% foetal bovine serum and 1% penicillin-streptomycin in a 37°C incubator with 5% CO_2. After removing the culture medium, they were incubated with 2.5 mM glucose and 1 mM $^{13}C_5$-stable isotope labelled glutamine in phosphate buffer for 18 h. Incubation was quenched by methanol, samples were silylated [*N*,*O*-bis(trimethylsilyl)trifluoroacetamide + 1% trimethylchlorosilane] and analysed by gas chromatography coupled to a triple quadrupole mass spectrometer (Scion, Brüker). Quantification of labelled versus natural isotope ions was performed in selected reaction monitoring mode.

Results

Identification of a mutation in *ALDH18A1* in Family FSP410

Under an autosomal dominant inheritance model, whole genome linkage analysis in Family FSP410 identified five putatively linked loci with multipoint logarithm of odds (LOD) scores reaching the maximal expected values for this pedigree (ranging from +1.64 to +1.95) as well as various uninformative regions with LOD scores varying from −1.90 to +0.70 (Supplementary Fig. 1). Whole exome sequencing performed in four patients provided 117 to 130 million reads per sample, 98% of which could be aligned to the targeted sequence. Mean depth of the targeted sequence was 125- to 132-fold. From 114 285 to 119 137 SNPs (single nucleotide polymorphisms) and from 10 546 to 11 159 indels were identified. Two missense variants respecting the abovementioned criteria segregated with the disease in the entire family: a c.359T > C/p.V120A

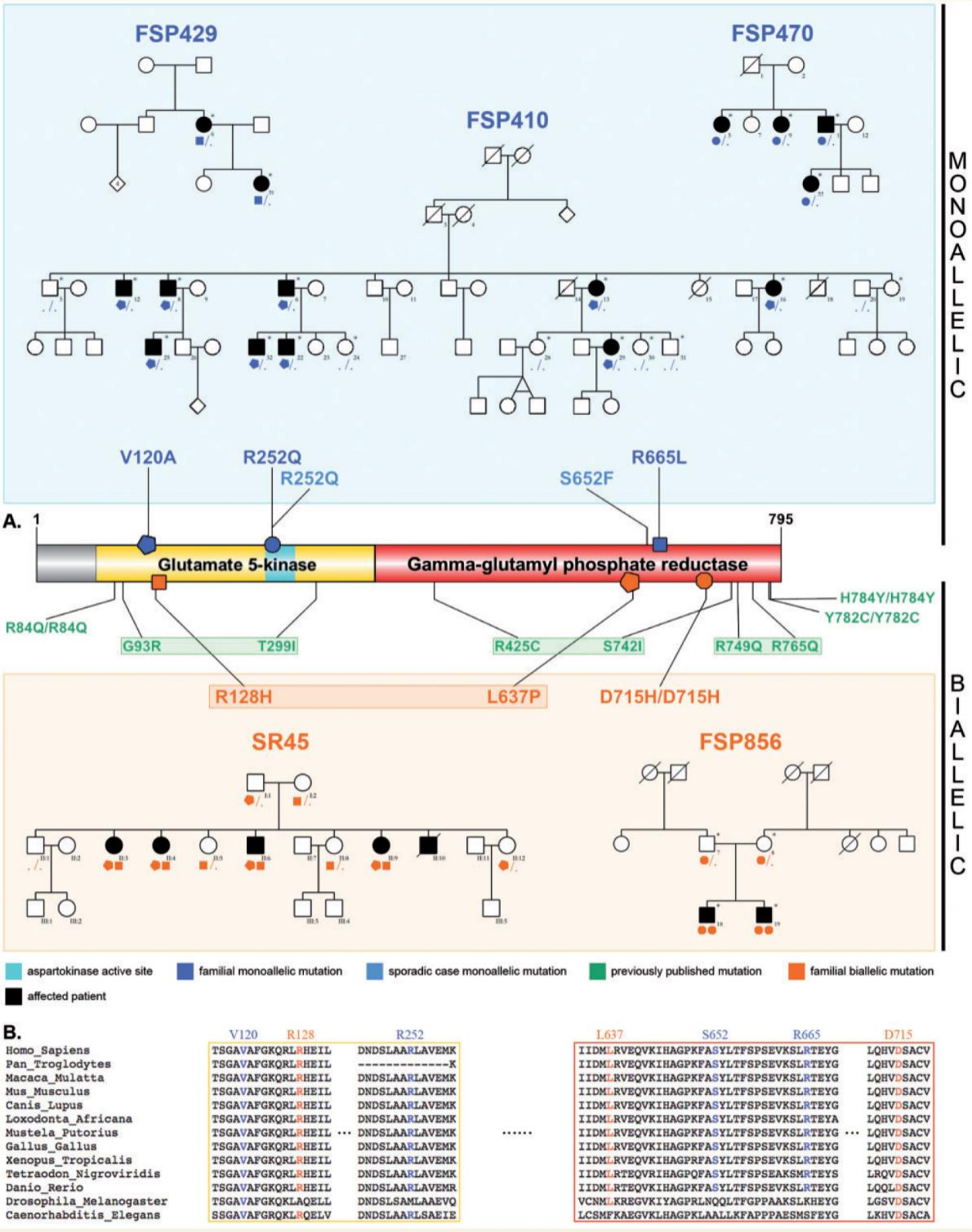


Figure 1 Pedigrees, *ALDH18A1* mutations and conservation across species. (**A**) Schematics of P5CS and its gamma-glutamyl kinase and gamma-glutamyl phosphate reductase domains. Newly described monoallelic mutations (blue), newly described biallelic mutations (orange), and previously reported biallelic mutations (green) are represented, along with pedigrees of the families of this study, showing the segregation. Sampled individuals are indicated with an asterisk. In autosomal dominant families, all affected patients carry one *ALDH18A1* mutation; the penetrance seems to be complete as no tested unaffected family member has any mutation. (**B**) Alignment of P5CS paralogues, showing the high conservation of all affected residues reported.

(NM_002860.3) change in *ALDH18A1* (OMIM 138250), at chr10:g.97 397 138; and a c.2250G > T/p.W750C (NM_001195263) change in *PDZD7* (OMIM 612971), at chr10:g.102 770 396. Only the *ALDH18A1* variant was predicted to be deleterious by all four prediction software used (Table 1) and was located in a functional domain of the protein. It affected a highly conserved amino acid across species (Fig. 1). *ALDH18A1* was previously implicated in an autosomal recessive syndrome that associates cutaneous, ocular, metabolic and neurological findings, with pyramidal signs (Baumgartner *et al.*, 2005), often limited to brisk reflexes; *PDZD7* contributes to Usher syndrome type II (OMIM 605472), either as a modifier or in a digenic pattern of transmission (Ebermann *et al.*, 2010). We found no additional causative *PDZD7* variants in 95 HSP index cases and had less *in silico* elements in favour of its pathogenicity; we therefore focused on *ALDH18A1*.

Additional families with *ALDH18A1* mutations

To confirm the implication of *ALDH18A1* in autosomal dominant HSP, and further assess its potential role in autosomal recessive HSP, we screened 530 patients through whole exome sequencing ($n = 435$) or panel sequencing ($n = 95$). Segregating heterozygous variants were found in two French families compatible with an autosomal dominant inheritance mode: a c.1994G > T/p.R665L change at g.97 371 129 in Family FSP429, and a c.755G > A/p.R252Q change at g.97 392 769 in Family FSP470. The same p.R252Q change was also identified in Patient 25014 with pure HSP living in the USA. Reference haplotypes for two SNPs flanking the *ALDH18A1* mutation (rs 9787589, rs 10882640) allowed excluding the existence of a common ancestor with Family FSP470, that presents homozygous variants at both positions. A heterozygous c.1949G > T/p.S652F change at g.97 371 198 was identified in a sporadic patient from Australia/UK (Patient GSHSP44). In two families with complex autosomal recessive HSP, biallelic variants were identified: compound heterozygosity for a c.1910T > C/p.L637P change at g.97 373 512 and a c.383G > A/p.R128H change at g.97 397 114 in Family SR45 from Portugal, and a homozygous c.2143G > C/p.D715H change at g.97 370 017 in Spanish Family FSP856. All mutations affected conserved amino acids and were absent or rare in exome databases (Fig. 1 and Table 1).

Clinical characteristics of patients with *ALDH18A1* mutations

In Family FSP410, clinical assessment was available for seven of nine patients (Table 2). The common clinical picture was a complex, slowly progressive but finally severe HSP, with onset in adolescence or adulthood (ranging from 14 to 59 years) and predominant gait signs, motor

Table 1 Characteristics of novel *ALDH18A1* mutations

Family	Mutation characteristics		Database frequency			Pathogenicity prediction					Conservation	
	hg19 (chr10)	cDNA and protein changes	dbSNP 138	EVS	ExAC	SIFT score	PolyPhen-2 HDIV	PolyPhen-2 HVAR	LRT score	MutationTaster score	GERP + + score	PhyloP score
FSP410	97397138	NM_002860.3: c.359T > C; p.Val120Ala	0	0	0/122000 +	0 (D)	1 (D)	0.995 (D)	D	D (0.999993)	5.6	2.134
FSP429	97371129	NM_002860.3: c.1994G > T; p.Arg665Leu	0	0	0/122000 +	0.06	0.949 (P)	0.78 (P)	D	D (0.999969)	5.6	2.639
FSP470	97392769	NM_002860.3: c.755G > A; p.Arg252Gln	0	0	0/122000 +	0.23	0.997 (D)	0.941 (D)	D	D (0.999779)	5.97	2.832
GSHSP44	97371168	NM_002860.3: c.1955C > T; p.Ser652Phe	0	0	0/122000 +	0.02 (D)	0.892 (P)	0.694 (P)	D	D (0.999952)	5.6	2.639
25014	97392769	NM_002860.3: c.755G > A; p.Arg252Gln	0	0	0/122000 +	0.23	0.997 (D)	0.941 (D)	D	D (0.999779)	5.97	2.832
FSP856	97370017	NM_002860.3: c.2143G > C; p.Asp715His	0	0	1/122892 - 0 hmz	0 (D)	1 (D)	1 (D)	D	D (0.999999)	5.43	2.722
SR45 – mutation 1	97397114	NM_002860.3: c.383G > A; p.Arg128His	0	0	4/122890 - 1 hmz	0.01 (D)	0.999 (D)	0.978 (D)	D	D (0.999981)	5.6	2.642
SR45 – mutation 2	97373512	NM_002860.3: c.1910T > C; p.Leu637Pro	0	0	0/122000 +	0 (D)	1 (D)	0.999 (D)	D	D (0.999999)	6.07	2.326

Summary of characteristics for all variants reported. The database frequencies were looked for in dbSNP138 (http://www.ncbi.nlm.nih.gov/SNP/snp_summary.cgi), Exome Variant Server (EVS, http://evs.gs.washington.edu/EVS) and Exome Aggregation Consortium (ExAC, http://exac.broadinstitute.org/). Pathogenicity scores are evaluated as follows: SIFT (http://sift.jcvi.org/) predicts deleteriousness (D) under 0.05; PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/) classifies SNPs as probably damaging (D; HDIV $\geq$ 0.957, HVAR $\geq$ 0.909), possibly damaging (P; 0.453 $\leq$ HDIV $\leq$ 0.956, 0.447 $\leq$ pp2_hdiv $\leq$ 0.908), or benign (B; HDIV $\leq$ 0.452, HVAR $\leq$ 0.446); LRT (Chun and Fay, 2009) differentiates variants with deleterious (D), neutral (N) or unknown (U) effect; MutationTaster (http://www.mutationtaster.org/) classifies them as 'disease_causing_automatic' (A), 'disease_causing' (D), 'polymorphism' (N) or 'polymorphism_automatic' (P) with a given probability value, 1 being the most probable. For both GERP + + (http://mendel.stanford.edu/SidowLab/downloads/gerp/) and PhyloP (http://compgen.bscb.cornell.edu/phast/help-pages/phyloPtxt), higher scores indicate better residue conservation.
Hmz = homozygous.

Table 2 Clinical characteristics of patients with *ALDH18A1* mutations: autosomal dominant families

Family N°/origin	FSP410/Italy							FSP429/France		FSP470/France				Spo/USA	Spo/Australia and UK
ALDH18A1 variant/ inheritance	p.Val120Ala/Autosomal dominant							p.Arg665Leu/Dominant		p.Arg252Gln/Dominant				p.Arg252Gln/ Sporadic	p.Ser652Phe/ Sporadic
Individual N° (sex)	29 (F)	13 (F)	6 (M)	32 (M)	8 (M)	16 (F)	22 (M)	6 (F)	21 (F)	5 (F)	9 (F)	11 (M)	22 (F)	25014 (F)	GSHSP44 (F)
Age at exam (years)	45	70	71	43	61	50	28	66	37	52	50	48	20	76	46
Age at onset (years)	18	22	49	14	59	42	?	24	13	19	21	43	1	31	44
Symptoms at onset	Muscle cramps	Weakness	Stiff legs	Stiff legs	Pain and weakness in LL	Stiff legs, mild instability	Reflex pyramidal syndrome	Stiff legs	Unsteadiness	Falls	Stiff legs	Stiff legs	Stiff legs	Falls	Toe walking, tripping
Disease duration (years)	27	48	22	29	2	25	?	34	24	33	29	5	19	45	12
Disability score (max 7)	4/7	5/7	7/7	3/7	NA	3/7	0/7	6/7	4/7	3/7	4/7	4/7	3/7	5/7	4/7
Spasticity at gait	Severe	Severe	Severe	Moderate	NA	Moderate	None	Severe	Severe	Moderate	Severe	Moderate	Moderate	Moderate	Mild
Spasticity at rest	Mild Ashworth 1/4	Mild	Severe	Mild	NA	Mild	None	None Ashworth 0/4	Moderate	Moderate Ashworth 2/4	Mild Ashworth 2/4	Mild	Mild	Severe	Mild
Weakness	Distal > proximal	Diffuse (3/5 MRC grading 4 limbs)	Proximal, distal	Distal > proximal (4 limbs)	L > R	Distal (3/5) > proximal	None	Distal	Distal > proximal	Proximal, distal	Proximal, distal	Proximal, distal	Proximal, distal	Distal > proximal	Distal > proximal
Increased reflexes LL	Yes (ankles abolished)	Yes (ankles abolished)	Yes (ankles abolished)	Yes (ankles abolished)	Yes	Yes	Yes	Yes (ankles decreased)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Increased reflexes UL	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Normal	Yes	Yes	Yes	Yes	Yes	Yes
Extensor plantar response	Indifferent	Yes	Yes	Indifferent	Yes (R)	Yes	R indifferent	Yes	Indifferent	Yes	Yes	Yes	Yes	Yes	Yes
Hoffman sign	NA	NA	NA	Yes	Yes	Yes	Yes	NA	Yes	Yes	Yes	Yes	No	NA	NA
Decreased vibration sense at ankles	ND	ND	Yes	Mild	NA	No	No	No	Mild	No	No	No	No	Moderate	No
Urinary symptoms	Pollakiury	No	Yes	Pollakiury	NA	Yes	No	No	Urinary incontinence	Urinary urgency	Urinary urgency	No	No	Urinary urgency	Urinary urgency
Cerebellar gait	No	No	Mild	Mild	NA	Mild	No	Yes (CCFS 0.939)	No	No	No	No	No	No	No
Cerebellar signs UL	No	No	Mild	No	NA	No	No	Yes	No	No	No	No	No	No	No
Dysarthria	Spastic	No	Anarthria (age 70)	Spastic	NA	Spastic	No	No	Yes	No	No	No	No	No	No
Cognitive impairment	No	No	Dementia (age 69)	No	Mild memory impairment	Mild memory impairment	No	No	No	No	No	No	No	No	No
Cutaneous findings	No	No	No	No	NA	No	No	No	No	No	No	No	No	Chronic cellulitis changes	No
Ocular findings	NA	NA	NA	NA	NA	NA	Congenital cataract	NA	NA	NA	NA	NA	Cataract (age 8 months)	Bilateral senile cataracts (69)	NA
Other clinical features	Pes cavus, neuropathic pain	Neuropathic pain	Low back pain, neuropathic pain (LL)	Pes cavus, infantile psychosis, mild horizontal nystagmus		Neuropathic pain (LL), gaze nystagmus, neurosensory deafness, temporal epilepsy	Mild pes cavus	Pes cavus	Pain (knee and back)	Pes cavus	Pes cavus, tremor (left UL)	Mitral leak	Gastrointestinal reflux	Resting tremor, uterine cancer (69)	Pes cavus, hip replacement, groin pain
MRI	Normal brain MRI (2002)	ND	ND	Mild CC atrophy	Normal brain MRI	Large cisterna magna on MRI	NA	White matter anomalies, pons hypersignal, thin dorsal cord	Thin dorsal cord	NA	Left arachnoidien cyst	NA	NA	Cortical atrophy (frontoparietal predominance), chronic microvascular disease (69)	Spinal cord atrophy
Other paraclinical tests	ENMG: motor neuronopathy (2009)	NA	NA	ENMG: motor neuronopathy (2014)	Normal ENMG	ENMG: motor neuronopathy (LL, L>R) (2004)	NA	NA	NA	NA	NA	NA	NA	NA	NA

CC = corpus callosum; CCFS = composite cerebellar functional severity score; ENMG = electro-neuro-myogram; HC = head circumference; ID = intellectual delay; L = left; LL = lower limbs; MD = motor delay; MRC = medical research council; NA = not available; ND = not done; PMD = psychomotor delay; R = right; UL = upper limbs.

neuropathy (3/4 tested) and spastic dysarthria (4/6). Mild cerebellar signs in gait (3/6) and upper limbs (1/6) were described; one patient presented with dementia at age 69. One carrier (Patient FSP410-22) presented with congenital cataract but also showed, on last examination at age 28, an isolated mild pyramidal syndrome of the four limbs with enhanced upper limbs reflexes, brisk patellar reflexes, absent ankle reflexes, bilateral Hoffman signs and unilaterally indifferent plantar cutaneous reflex with no spasticity. Families FSP429 and FSP470 presented with moderate to severe spasticity in gait with few or no associated signs. Cerebellar signs were present in Family FSP429 (1/2), cataracts and gastrointestinal reflux in Patient FSP470-22. Brain MRI was abnormal in Patient FSP429-6 with mild FLAIR hyperintensity of the corticospinal tract posterior of the ventricles, a localized and heterogeneous T_2 hyperintensity in the pons, and moderate spinal cord atrophy (Supplementary Fig. 2). Both additional heterozygous cases presented with pure HSP, with spinal cord atrophy on MRI in Patient GSHSP44. Pes cavus was reported in 7 of 15 patients with monoallelic *ALDH18A1* variant.

In patients from autosomal recessive HSP Families FSP856 and SR45 (Table 3), cognitive impairment was primary, together with moderate to severe early onset HSP. Three of four patients in Family SR45 had a continuum of presentations from tetraparesis to tetraplegia, and two showed pseudobulbar palsy. Both brothers from Family FSP856 had upper limb postural tremor suggesting cerebellar involvement. They also presented microcephaly and facial dysmorphism (2/2), while patients in Family SR45 had short stature (4/4), cataract (1/4) and facial dysmorphism (3/4). Brain MRI showed a thin corpus callosum in patients from Family SR45 but was normal in one patient from Family FSP856 (Supplementary Fig. 2). Of note, both the father and the mother in Family FSP856 carry the heterozygous mutation, while none of them show clinical signs on examination.

Biochemical analyses in blood and dermal fibroblasts

Blood amino acid chromatography (Table 4) was performed in four patients from two unrelated autosomal dominant families. Three family members of Family FSP410 had low citrulline levels, associated with low proline (2/3), ornithine (1/3, 1/3 borderline), and borderline levels of arginine (3/3). A profound decrease of plasma citrulline was also observed in Patient FSP429-21, associated with low arginine and borderline ornithine levels. By contrast, amino acid levels were normal in Patient FSP856-18 with autosomal recessive inheritance (data not shown). In all tested patients, levels of other amino acids in blood and all amino acids in urine were within normal range. The amino acids profiles from the patients in this study, along with the two first patients reported with *ALDH18A1* mutations (Baumgartner *et al.*, 2005), were compared to 5023 profiles from Necker Hospital full cohort of patients aged over 15 years that had no known metabolic disorders. The comparison showed that low levels of ornithine, citrulline, arginine and proline (Fig. 3A) were significantly predictive of *ALDH18A1* mutations (area under the receiver operating characteristic curve 97%, with 95% confidence interval: 96–99%, Supplementary Fig. 3A). Adding elevated glutamine as a covariate may increase prediction (Supplementary Fig. 3B).

In cultured dermal fibroblasts from two patients of Family FSP410, proline biosynthesis was estimated by the stable isotope label enrichment in proline after loading for $^{13}C_5$-labelled glutamine (Baumgartner *et al.*, 2005). Assuming a stationary phase at 18-h incubation, the two patients harbouring *ALDH18A1* mutations had 42% residual flux (stable isotope ratios 0.68 + 0.08 and 0.95 + 0.06, average: 0.81) compared to age-matched healthy controls (stable isotope ratio 1.83 ± 0.05 and 2.01 ± 0.14, average: 1.92) (Fig. 3B). This confirms a reduction of proline biosynthesis consistent with an enzymatic deficiency of P5CS.

Discussion

Here we report seven pedigrees with *ALDH18A1* mutations. Two families with biallelic mutations presented complex autosomal recessive HSP with marked cognitive impairment and no cutaneous involvement. Remarkably, we also identified, in three independent families, monoallelic *ALDH18A1* mutations segregating with autosomal dominant HSP, together with low plasma levels of specific amino acids, especially citrulline, and *ex vivo* demonstration of a metabolic block at the level of P5CS. Two additional sporadic cases harboured monoallelic mutations in *ALDH18A1*. Therefore, this study broadens the clinical spectrum of autosomal recessive *ALDH18A1*-linked pathology, and describes a new mode of inheritance associated with a potential trait biomarker in plasma. Expanding phenotype and describing new modes of transmission becomes more and more frequent in HSP (Esteves *et al.*, 2014; Tesson *et al.*, 2015); such bimodal transmission has also been described long ago, for example in human myotonia, caused by either dominant or recessive mutations of *CLCN1* (Koch *et al.*, 1992).

ALDH18A1 encodes P5CS, a bifunctional enzyme that catalyses the first common step in proline and ornithine biosynthesis (Hu *et al.*, 1999, 2008*b*). This enzyme, located in the mitochondrial inner membrane, catalyses the conversion of glutamate to gamma-glutamyl semi-aldehyde (Fig. 3), which spontaneously converts to pyrroline-5-carboxylate (P5C). P5C can thereafter be metabolized to proline by P5C reductase (PYCR1), or to ornithine, citrulline and finally arginine through the urea cycle (Fig. 3). P5CS comprises two domains, with different enzymatic activities: an N-terminal ATP-dependant gamma-glutamyl kinase domain, responsible for the

Table 3 Clinical characteristics of patients with *ALDH18A1* mutations: autosomal recessive families

Family N°/Origin	FSP856/Spain		SR45/Portugal			
ALDH18A1 variant/ Inheritance	p.Asp715His, p.Asp715His/Recessive		p.Arg128His, p.Leu637Pro/Recessive			
Individual N° (sex)	18 (M)	19 (M)	II:3 (F)	II:4 (F)	II:6 (M)	II:9 (F)
Age at exam (years)	42	39	44	49	46	43
Age at onset (Years)	7	7	<1	<1	<1	<1
Symptoms at onset	Toe walking, ID, speech delay	ID, gait difficulties	MD, growth retardation	MD, growth retardation	PMD	Severe MD, growth retardation
Disease duration (years)	35	32	44	49	46	43
Disability score (max 7)	6/7	3/7	7/7	7/7	6/7	7/7
Spasticity at gait	Severe	Moderate	Severe, ambulation lost at 30	Severe	Severe, ambulation lost at 15	Severe, no ambulation
Spasticity at rest	Severe	Moderate	Severe Ashworth 4/4	Severe Ashworth 3/4 UL; 4/4 LL	Severe Ashworth 4/4 LL; 2/4 UL	Severe Ashworth 4/4
Weakness	Proximal, distal	No	Tetraplegia, drop-feet	Tetraparesia, drop-feet	MRC grading LL 0/5, UL 4/5	Tetraplegia
Increased reflexes LL	Yes	Yes	Yes (ankles abolished)	Yes (ankles abolished)	Yes (ankles abolished)	Not awakened
Increased reflexes UL	Yes	Yes	Yes	Yes	Yes	Not awakened
Extensor plantar response	Yes	Yes	Yes	Yes	Yes	Indifferent
Hoffman sign	NA	NA	NA	NA	NA	NA
Decreased vibration sense at ankles	Mild	Mild	NA	NA	No	NA
Urinary symptoms	Urinary retention, pollakiuria	No	Incontinence	Incontinence	Incontinence	Incontinence
Cerebellar gait	NA	No	No	No	No	NA
Cerebellar signs UL	Postural tremor	Postural tremor	No	No	No	NA
Dysarthria	No	No	Pseudobulbar, loss of speech (age 30)	Pseudobulbar	Spastic	No speech ever
Cognitive impairment	Intellectual deficit IQ 49	Intellectual deficit IQ 35-50	ID, progressive deterioration	ID, progressive deterioration	ID, progressive deterioration	Profound
Cutaneous findings	No	No	No	No	No	No
Ocular findings	NA	NA	NA	NA	Probable cataract (on neurological examination)	NA
Other clinical features	Microcephaly (HC 52 cm, −2 SD), facial dysmorphy	Microcephaly (HC 52 cm, −2 SD), facial dysmorphy	Microcephaly, growth retardation, facial dysmorphy; generalized muscle atrophy	Microcephaly, short stature, generalized muscle atrophy	Mild facial dysmorphy, distal amyotrophy (UL, LL)	Severe growth retardation, cyphoscoliosis, archaic reflexes, extreme muscle atrophy
MRI	Normal brain MRI	NA	NA	NA	CC atrophy, periventricular white matter anomalies, mild cortical atrophy	NA

CC = corpus callosum; CCFS = composite cerebellar functional severity score; ENMG = electro-neuro-myogram; HC = head circumference; ID = intellectual delay; L = left; LL = lower limbs; MD = motor delay; MRC = medical research council; NA = not available; PMD = psychomotor delay; R = right; UL = upper limbs.

Table 4 Amino acid levels on plasma chromatography

Patient	Age	Proline	Ornithine	Citrulline	Arginine	Valine	Leucine	Isoleucine	Threonine	Glutamine	Alanine	Methionine	Lysine
FSP410-032	43	163	**35**	**12**	*58*	*190*	116	*51*	*69*	517	*237*	22	*146*
FSP410-029	45	**117**	83	**8**	*63*	215	110	61	208	740	400	23	201
FSP410-013	70	**73**	*46*	**8**	*54*	*178*	*108*	*51*	*91*	632	*277*	*18*	178
FSP429-021	40	209	*54*	**1**	**28**	267	119	*52*	114	470	304	*19*	175
Normal ranges		150–220	50–100	20–35	60–100	210–280	100–150	50–80	95–195	430–670	285–415	20–35	155–230

Values are given in µmol/l. Bold indicates low values, italics indicates borderline values.

glutamate phosphorylation to gamma-glutamyl phosphate, and a C-terminal NADPH-dependant gamma-glutamyl phosphate reductase domain, which catalyses the reduction and conversion to gamma-glutamyl semi-aldehyde (Fichman *et al.*, 2014). Although P5CS expression in the brain is not strong (Hu *et al.*, 1999), it has a measurable activity (Hu *et al.*, 2008*b*).

ALDH18A1 mutations account for a phenotypic spectrum ranging from pure autosomal dominant to complex autosomal recessive HSP

In humans, mutations in two genes coding for enzymes involved in proline biosynthesis have been described in overlapping neurocutaneous syndromes; *ALDH18A1* and *PYCR1*. *PYCR1* mutations are implicated in an autosomal recessive neurocutaneous syndrome with mental retardation, cutis laxa, joint hyperlaxity, progeroid dysmorphy, growth retardation, microcephaly, and corpus callosum dysgenesis (Reversade *et al.*, 2009; Dimopoulou *et al.*, 2013; Gardeitchik *et al.*, 2014). These patients present normal amino acid levels (Reversade *et al.*, 2009).

As for *ALDH18A1*, the first two patients with demonstrated P5CS deficiency carried a homozygous missense mutation in the glutamate-5-kinase domain of the protein (Baumgartner *et al.*, 2000, 2005), and presented a syndrome associating cutis laxa, joint hyperlaxity, hypotonia, developmental delay, cataract, pyramidal signs and plasma amino acids anomalies—low proline, ornithine, citrulline and arginine. Subsequently, Bicknell *et al.* (2008) reported four siblings with a homozygous missense mutation in the gamma-glutamyl phosphate reductase domain of P5CS; the clinical picture was very similar with the exception of plasma amino acids levels, which showed no abnormalities. Until now, 15 patients with *ALDH18A1* biallelic mutations have been reported (Baumgartner *et al.*, 2000, 2005; Bicknell *et al.*, 2008; Skidmore *et al.*, 2011; Martinelli *et al.*, 2012; Zampatti *et al.*, 2012; Fischer *et al.*, 2014; Gardeitchik *et al.*, 2014; Handley *et al.*, 2014; Wolthuis *et al.*, 2014). Developmental delay, failure to thrive, hypotonia, cognitive impairment, cutis laxa, joint hyperlaxity, growth retardation and cataract are cardinal signs (Table 5). Dystonia, seizures in infancy, thin corpus callosum, pes planus, hip dislocation, bone malformations, gastro-oesophageal reflux and inguinal hernia are frequent signs orientating towards the molecular diagnosis. Pyramidal signs are often reported (12/15), but not cardinal; most of the patients presented brisk reflexes and pyramidal syndrome was judged severe in three only.

The patients that we report with biallelic *ALDH18A1* mutations (Families FSP856 and SR45) share common features with those previously reported suffering from neurocutaneous syndromes: developmental delay (4/6), intellectual deficiency (6/6), short stature (4/6), facial dysmorphism (6/6), cataract (1/6) and thin corpus callosum (1/3). Of interest is the severity of pyramidal signs, with moderate to severe spasticity (6/6), to tetraplegia (1/6) and pseudobulbar palsy (2/6). Although severe HSP was previously reported in *ALDH18A1* patients, its frequency was low (3/15). Moreover, though cutis laxa has been reported to improve or disappear with age in some patients (Bicknell *et al.*, 2008), it was never noted in any of our patients, while it was previously considered as an obligate sign. We therefore are expanding the clinical spectrum of biallelic *ALDH18A1* mutations, with emphasis on the pyramidal component alteration; and are establishing that the cutaneous abnormalities may be lacking.

More importantly, we also report the first three families with autosomal dominant HSP segregating *ALDH18A1* mutations. The following arguments are in favour of the pathogenicity of the variants we report: (i) the absence of reported mutations in all examined public databases (~63 000 exomes); (ii) amino acid conservation through phylogenesis; (iii) *in silico* pathogenicity predictions (Table 1); (iv) location within a putatively linked locus in Family FSP410 (Supplementary Fig. 1); (v) absence of pathogenic variants in all known HSP genes; (vi) multiplicity of the mutated families and multiplicity of patients with the mutations amongst the families; (vii) location within the aspartokinase active site for p.R252Q; (viii) a second occurrence of the p.R252Q mutation with a different haplotype; (ix) abnormal plasma amino acids levels in carriers from two unrelated families (Families FSP410 and FSP429); and (x) reduced glutamate-to-proline metabolic flux in dermal fibroblasts of two patients of Family FSP410 (42% of levels measured in control fibroblasts). However, the clinical picture of these patients largely differs from biallelic mutation carriers. Spasticity is the cardinal

Table 5 Summary of clinical characteristics of published families with *ALDH18A1* mutations

Reference	Baumgartner *et al.*, 2000, 2005	Bicknell *et al.*, 2008	Skidmore *et al.*, 2011	Martinelli *et al.*, 2012	Zampatti *et al.*, 2012; Gardeitchik *et al.*, 2014	Wolthuis *et al.*, 2014	Fischer *et al.*, 2014	Fischer *et al.*, 2014	Handley *et al.*, 2014	This study (FSP856)	This study (SR45)	This study (FSP410)	This study (FSP429)	This study (FSP470)	This study (GSHSP44)	This study (25014)
Mutations (protein level)	R84Q/R84Q	H784Y/H784Y	c.1923 + 1 G > A/c. 1923 + 1 G > A	G93R/T299I	S742I/R425C	Y782C/Y782C	V601Gfs*/ V601Gfs*	L711Cfs*/ L711Cfs*	R749Q/R765Q	D715H/ D715H	R128H/L637P	V120A	R665L	R252Q	S652F	R252Q
Functional effect	Reduced P5CS activity	Normal PRO and ORN synthesis	Reduced fibroblasts proliferation and P5CS expression	Decreased P5CS activity, dissociation to dimers or monomers; no mitochondrial alterations in cultured fibroblasts			Abolished P5CS expression, swollen mitochondria in fibroblasts (electron microscopy)		Reduced P5CS expression, lipid droplet enlargement	NA	NA	Lowered PRO synthesis in fibroblasts	NA	NA	NA	NA
Low plasma amino acids levels (other findings)	ORN, CITRU, ARG, PRO (paradoxal hyperammonaemia)	Normal (mild post-prandial hypoornithinaemia in 2/3)	Normal	ORN, CITRU, ARG, PRO (paradoxal hyperammonaemia)	Normal	Normal	CITRU, ARG	NA	Normal	NA	NA	CITRU (3/3), PRO (2/3), low or borderline ORN(2/3), and ARG (3/3)	CITRU, ARG, (borderline ORN)	NA	NA	NA
Neurological signs at onset	DD, hypotonia	DD, hypotonia	DD, hypotonia	DD, hypotonia	DD, hypotonia	DD, hypotonia	DD, hypotonia		DD, hypotonia	Intellectual deficiency	Global delay, growth retardation	Pyramidal signs	Pyramidal signs	Pyramidal signs	Pyramidal signs	Pyramidal signs
Pyramidal signs	Yes (2/2), severe in 1	Hypertonicity LL (4/4), UL (2/4)	Yes	NA	Yes	Brisk reflexes	NA	NA	Brisk reflexes	Moderate to severe	Severe, tetraplegia (2/4) or tetraparesis (1/4)	Moderate to severe	Severe	Moderate to severe	Mild	Moderate to severe
Other neurological findings	Peripheral axonal neuropathy, seizures in infancy (1/2), distal dystonia (1/2)	Seizures in infancy (2/4), absent speech	Flexion contractures of elbows, wrists (32w gestation)	Seizures in infancy	Seizures, dystonia, absent speech		Disseminated tremors, finger contractures	Seizures	Joint contractures	No	No	Mild cerebellar signs (2/6), motor neuropathy (2/6), dementia (1/6)	Cerebellar signs (1/2)	No	No	No
Cutis laxa (other cutaneous findings)	Yes	Yes	Yes (sparse hair)	Yes	Yes (sparse hair)	Yes	Yes	Yes	Yes	No	No	No	No	No	No	Chronic cellulitis changes
Skeletal findings	Joint hyperlaxity, pes planus	Joint hyperlaxity, hip dislocation (3/4), kyphoscoliosis	NA	Joint hyperlaxity, pes planus, coxa valga	Joint hypermobility, osteopenia, rib anomalies	Joint laxity, bilateral hip dislocation and dysplasia	Unilateral hip dislocation, kyphosis	Joint hypermobility, congenital luxation of left hip	Bilateral coxa valga			Pes cavus (3/6)	Pes cavus (1/2)	Pes cavus (2/4)	Pes cavus, hip replacement	
Facial dysmorphy	Yes	NA	Yes	Yes	Yes	NA	Yes	NA	Yes	Yes	Yes	No	No	No	No	No
Growth retardation/ microcephaly	Yes/yes	Yes/yes	Yes/yes	Yes/macrocephaly	Yes/no	Yes/yes	Yes/yes	Yes/yes	Yes/yes	No/yes	Yes/yes	No/no	No/no	No/no	No/no	No
Cataracts (other ocular findings)	Yes	Yes 1/4	(Corneal clouding)	Yes	Yes	(Retinitis pigmentosa)	Yes	(Corneal clouding)	Yes 1/2	NA	Yes 1/4	Yes 1/6	NA	Yes 1/4	NA	No
Gastro-oesophagal reflux	2/2	NA	NA	NA	Yes	NA	NA	NA	NA	NA	NA	NA	NA	1/4	NA	NA

(continued)

Table 5 Continued

Reference	Baumgartner *et al.*, 2000, 2005	Bicknell *et al.*, 2008	Skidmore *et al.*, 2011	Martinelli *et al.*, 2012	Zampatti *et al.*, 2012; Gardeitchik *et al.*, 2014	Wolthuis *et al.*, 2014	Fischer *et al.*, 2014	Fischer *et al.*, 2014	Handley *et al.*, 2014	This study (FSP856)	This study (SR45)	This study (FSP410)	This study (FSP429)	This study (FSP470)	This study (GSH5SP44)	This study (25014)
Inguinal hernia	NA	1/4	Bilateral	Bilateral	1/2	Yes	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Vascular findings	NA	NA	Arterial tortuosities	NA	NA	NA	Cardiovascular malformations	Narrow aortic arch, probable aneurysm	NA	NA	NA	NA	NA	Mitral leak (1/4)	NA	NA
Other						Abnormal fat distribution, undescended testis	Bilateral hearing loss		Undescended testis			Infantile psychosis (1/6)				
MRI	NA	Paucity of WM, thin CC	NA	Mild cortical atrophy, thin CC, tortuosity of brain vessels	Cerebellar vermis hypoplasia, thin CC	Thin CC	Normal	Agenesis of CC	Thin CC, prominent lateral and 3rd ventricles, cerebellar hypoplasia, mildly delayed myelination	Normal	Thin CC, mild cerebellar atrophy, WM anomalies	Mild CC atrophy (1/3)	WM anomalies, thin dorsal cord, pons hypersignal (1/2)	Left arachnoidien cyst (1/4)	Spinal cord atrophy	Cortical atrophy (frontoparietal predominance)

ARG = arginine; CC = corpus callosum; CITRU = citrulline; DD = developmental delay; LL = lower limbs; NA = not available; ORN = ornithine; PRO = proline; UL = upper limbs; WM = white matter.

and first presentation, and is often pure with few or no accompanying signs. Seven of 15 patients had pes cavus that could be a consequence of the pyramidal syndrome, while pes planus was classically described in autosomal recessive pathology. Interestingly, Patients FSP410-22 and FSP470-22 presented with congenital cataract, which is a distinctive sign, when present, in patients with biallelic *ALDH18A1* mutations (Baumgartner *et al.*, 2000, 2005; Zampatti *et al.*, 2012). A peripheral neuropathy was demonstrated in three of four patients that had electromyography in Family FSP410, but was clinically present in 9 of 21 patients, as in the first reported patients with biallelic *ALDH18A1* mutations (Baumgartner *et al.*, 2005). Cerebellar signs were noted in 4 of 14 autosomal dominant patients, as in two of six patients with biallelic mutations. It has never been reported before, but cerebellar atrophy was evidenced in several reports (Zampatti *et al.*, 2012; Handley *et al.*, 2014). These elements are again in favour of the implication of *ALDH18A1* mutations in these cases. Of note, the clinical presentation was not homogeneous among the respective families, especially for Family FSP410. Age, disease duration, lifestyle, and dietary factors are all potential modifiers underlying this variability.

What is the pathogenic mechanism linked to *ALDH18A1* mutations?

In autosomal recessive families, all previously reported patients had either homozygous variants or compound heterozygous mutations affecting the same functional domain and, therefore, a presumably complete enzymatic block of one of the P5CS-catalysed reactions (Fig. 3). In addition, abnormal metabolic profiles were only reported in patients with mutations affecting the glutamate-5-kinase domain (Baumgartner *et al.*, 2000; Martinelli *et al.*, 2012) or with a complete abolition of protein expression (Fischer *et al.*, 2014). Consistently, the amino acid profile is normal in Patient FSP856-18 that harbours mutations that do not affect the glutamate-5-kinase domain. Of note, Family SR45 is the first reported with compound heterozygous missense mutations involving both domains of P5CS; here, a complete enzymatic block is not expected.

In autosomal dominant families, the most striking anomalies are the low to very low citrulline levels in the plasma of all patients tested, regardless of the protein domain affected by the mutation, together with a less consistent, decreased ornithine, arginine and proline levels. The sum of age-normalized values of these four amino acids is obviously reduced in our patients, as well as in the first family in which the disorder was identified (Fig. 2A), and is significantly, though incompletely, predictive of *ALDH18A1* mutations (Supplementary Fig. 3A). Dermal fibroblasts from Family FSP410 showed a decrease in proline biosynthesis by a stable isotope label substrate-loading test (Fig. 2B). The degree of observed residual flux as measured by the accumulation of labelled proline relative to the

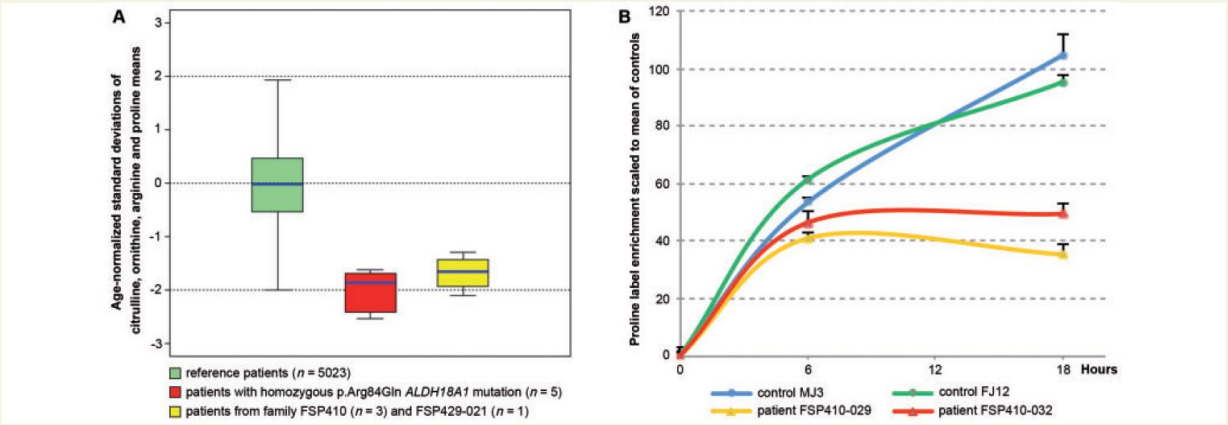


Figure 2 Plasma amino acid profiles and proline biosynthesis flux analysis. (**A**) Boxplots show the mean of four plasma amino acid levels (citrulline, ornithine, proline and arginine) expressed as their age-normalized standard deviations (*y*-axis; see 'Materials and methods' section) compared among three groups of patients: the full Necker hospital cohort of patients with no known metabolic disorders, aged >15 years (n = 5023; green); patients from the first family in which *ALDH18A1* mutations were reported (n = 5; red); and the patients from Families FSP410 and FSP429 (this study) for which plasma amino acids were available (n = 4; yellow). The average of the four amino acid levels in *ALDH18A1* mutated patients maps at approximately −2 SD. (**B**) Time course of stable isotope ratio (labelled versus natural) enrichment of proline reflecting the rate of proline biosynthesis from $^{13}C_5$-labelled glutamine in fibroblasts. Two *ALDH18A1*-mutated patients (red and orange lines) are compared to two controls, i.e. members of the same family with no mutations (green and blue lines). *Y*-axis: label enrichment scaled to the mean of the controls at 18-h incubation. *X*-axis: hours of incubation in the presence of the tracer. Assuming a stationary phase at 18-h incubation time, proline biosynthesis in two patients harbouring *ALDH18A1* p.Val120Ala mutation was estimated to be 42% of controls.

natural isotope is comparable to that measured in the first two patients reported to have biallelic *ALDH18A1* mutations (40%; Baumgartner *et al.*, 2005). However, the method we used differs from that report, and a direct comparison of the effect of different mutations needs to be performed for final conclusion. Regarding the pathogenic mechanisms in autosomal dominant patients, investigations are needed to directly compare the effects of homozygous and heterozygous mutations, especially to evaluate whether the latter leads to haploinsufficiency or a dominant negative gain-of-function. A previous report on a patient homozygous for a frameshift mutation resulting in no protein expression (Fischer *et al.*, 2014) argues against haploinsufficiency as the pathogenic mechanism underlying dominant mutations.

Furthermore, for autosomal recessive patients for which the amino acid profile was normal, another moonlighting, unknown, function of P5CS might be involved in the pathophysiology. Proline metabolism is a major actor in adaptation to osmotic stress and cellular redox stress (Hu *et al.*, 2008*a*, *b*; Yasuda *et al.*, 2013; Fichman *et al.*, 2014). It is also involved in the apoptosis pathway; of note P5CS was shown to be upregulated by p53 (Hu *et al.*, 2008*b*). Another field of interest is the mitochondria, as defects in mitochondrial structure have been described in *PYCR1*-linked neurocutaneous disease (Reversade *et al.*, 2009). No such abnormalities had previously been identified in P5CS deficiency syndromes (Skidmore *et al.*, 2011). Recently, Fischer *et al.* (2014) described swollen mitochondria in fibroblasts from their P5CS-deficient patient by electron microscopy on skin biopsies, but they could not observe them on cultured fibroblasts. In cultured fibroblasts from Patients FSP410-29 and FSP410-32, we found normal mitochondrial structure or membrane potential (data not shown). Renewed efforts are warranted to elucidate the cellular basis of *ALDH18A1*-related phenotypes.

Mitochondrial ornithine deficiency as a possible common scheme for multiple metabolic disorders

There are few metabolic causes of persistent hypocitrullinaemia after ruling out obvious and usually transient mechanisms of intestinal dysfunction: (i) urea cycle defects—notably, ornithine transcarbamylase (OTC), carbamoylphosphate synthase 1 (CPS1) and *N*-acetylglutamate synthetase (NAGS) deficiencies; (ii) ornithine translocase deficiency (or triple H syndrome); (iii) respiratory chain deficiencies (Rabier *et al.*, 1998); and (iv) P5CS/*ALDH18A1* deficiency (Baumgartner *et al.*, 2000, 2005). Several biochemical hints allow for distinction between the enzymatic defects. Likewise, proline is decreased in P5CS deficiency only, orotic aciduria is present in OTC deficiency only and hyperammonemia is mainly observed in the fed state for urea cycle defects, whereas it manifests during fasting in P5CS deficiency. Decreased plasma citrulline levels could therefore be an excellent trait biomarker for *ALDH18A1*-linked autosomal dominant HSP, particularly as the clinical picture is not specific (relatively pure HSP form). Furthermore, the combination of low ornithine, citrulline, arginine and proline

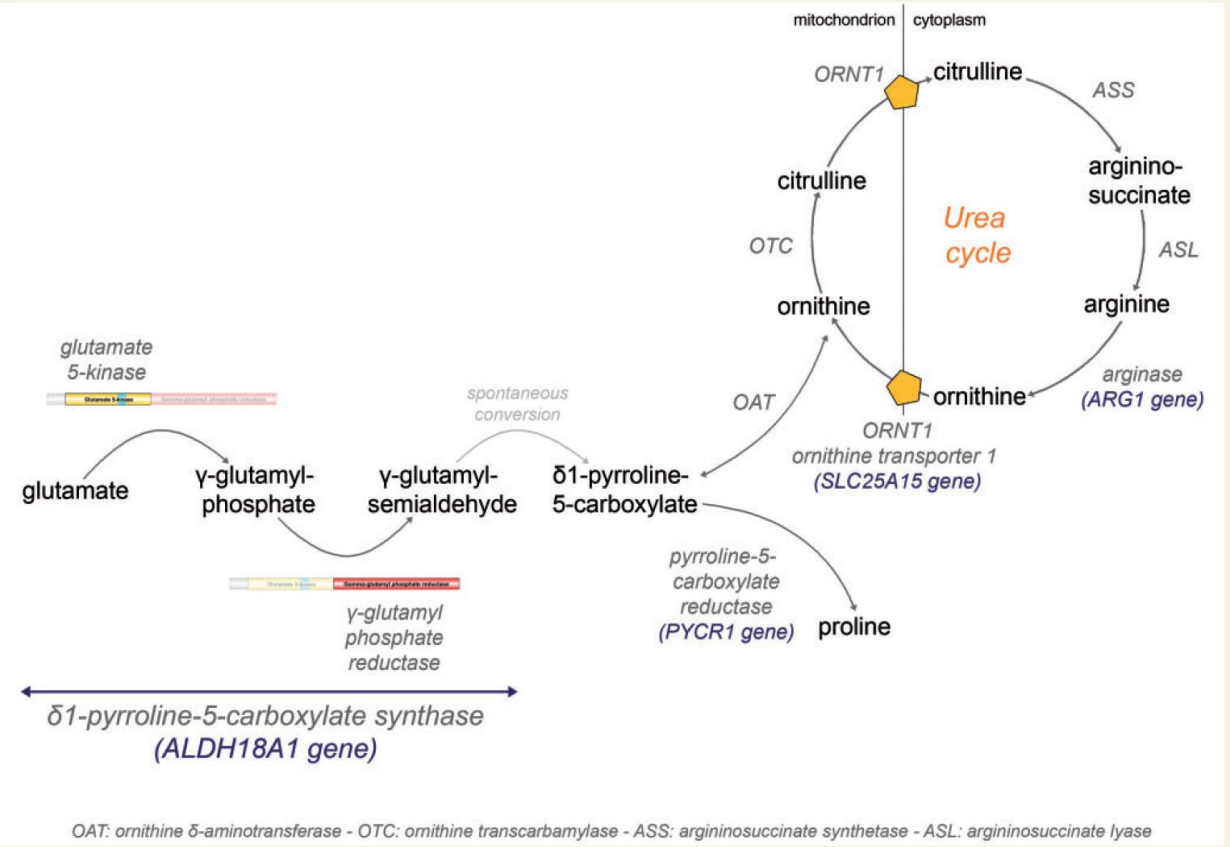


Figure 3 Delta-1-pyrroline-5-carboxylate synthase (P5CS) pathway. Schematic of the P5CS pathway. Enzyme-catalysed steps of proline and ornithine synthesis from glutamate, as well as urea cycle, are in dark grey. Mutations in *ALDH18A1* and *PYCR1* were previously reported in autosomal recessive neurocutaneous syndromes (Baumgartner *et al.*, 2000; Reversade *et al.*, 2009), with some pyramidal involvement. Deficiencies for arginase (*ARG1* mutations) and ornithine transporter (*SLC25A15* mutations, triple H syndrome) also share features of motor neuron degeneration (Sedel *et al.*, 2007). We speculate that a decrease in the mitochondrial ornithine pool may be the common feature of the latter two and *ALDH18A1*-linked autosomal dominant HSP.

in plasma is significantly predictive of *ALDH18A1*-linked autosomal dominant HSP (Supplementary Fig. 3A).

The fact that the urea cycle defects and autosomal recessive HSP share common biomarkers as well as features of motor neuron degeneration is of interest and suggests that common metabolic mechanisms may be involved. Arginase deficiency (*ARG1* mutations), an enzymatic step of the cytosolic part of the urea cycle (Fig. 3), is characterized by very high plasma levels of arginine and presumably decreased ornithine recycling. Triple H syndrome (*SLC25A15* mutations), a defect of the translocase responsible for transporting ornithine from the cytosol into the mitochondria (Fig. 3), leads to high plasma levels of ornithine, thus reflecting low mitochondrial ornithine. Plasma citrulline is normal in arginase deficiency, but decreased in triple H syndrome. Both metabolic diseases are associated with HSP in adults (Sedel *et al.*, 2007). Therefore, while plasma citrulline seems to be the most consistent biomarker of *ALDH18A1*-linked autosomal dominant HSP, a decrease in the mitochondrial pool of ornithine may be the common feature of arginase deficiency, triple H syndrome and P5CS deficiency (Fig. 3). We speculate that mitochondrial ornithine deficiency may be responsible for motor neuron degeneration in these diseases. This may have important therapeutic implications for patients with *ALDH18A1*-linked autosomal dominant HSP, because long-term supplementation in citrulline should be considered by analogy with its successful usage in patients with inborn errors of metabolism that directly or indirectly affect the urea cycle function—i.e. OTC deficiency or lysinuric protein intolerance (Haberle *et al.*, 2012).

In conclusion, we describe seven novel mutations in *ALDH18A1* in seven HSP pedigrees. Six patients from two families with complex autosomal recessive HSP broaden the phenotypic spectrum linked to this gene, with an emphasis on the pyramidal component and the absence of cutaneous signs. More importantly, we report

ALDH18A1 mutations segregating in an autosomal dominant pattern in 17 patients from five pedigrees with pure or slightly complex HSP, thus adding a new gene involved in autosomal dominant HSP. We also describe the first potential plasma biomarker of autosomal dominant HSP that can be implemented in daily practice. Due to the great diversity of HSP, plasma biomarkers are invaluable diagnostic tools for the screening of certain genetic entities among all possible aetiologies of spastic paraplegia or to validate the causality of novel variants found in high throughput sequencing methods in genetic diagnosis. Likewise, plasma very long chain fatty acids and 25-/27-hydroxycholesterols are now established biomarkers of HSP, leading to more effective detection and dedicated care of adrenomyeloneuropathy (*ABCD1* mutations) and SPG5 (*CYP7B1* mutations) (Schule *et al.*, 2010; Kemp *et al.*, 2012). Provided that our biochemical findings are validated in a larger population of *ALDH18A1*-linked autosomal dominant HSP patients, we recommend the use of plasma amino acid chromatography as part of diagnostic work-up in HSP. These results are also illustrative of the notion of 'genetics remodelling' brought by Next Generation Sequencing, as has been demonstrated for other genes (Esteves *et al.*, 2014; Synofzik *et al.*, 2014; Tesson *et al.*, 2015). The questioning of assumed transmission modes and new delineations in phenotype-genotype correlations are hallmarks of that field. Finally, we confirm that the metabolism of amino acids is an important cellular pathway in HSPs, which may help in the understanding of their physiopathology.

Acknowledgements

The authors are grateful to the patients, to Prs G. Ponsot, A. David, J-Ph. Neau, V. Paquis and Drs M. Clanet, P. Heine, A. Jacquette, L. Freemann, R. De Paz, M.L Monin and M. Aupy, for referring patients, to Justine Guégan, Dr C. Nava and Dr I. Moszer for Bioinformatic analysis, to Wassila Carpentier for SNP genotyping, to Sarah Boster for manuscript proofreading, to the Genotyping and Sequencing facility and the DNA and cell bank of the ICM for sample preparation and processing. We thank the Centre de Référence des Maladies Métaboliques (Necker Hospital, Paris) for sharing biochemical data of their cohort of patients.

Funding

The authors' work was partially funded by the Agence Nationale de la Recherche (SPATAX-QUEST, to G.S.), the Fundação para a Ciência e a Tecnologia (FCT-ANR/BEX-GMG/0008/2013, to I.A.), the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 2012-305121 'Integrated European –omics research project for diagnosis and therapy in rare neuromuscular and neurodegenerative diseases (NEUROMICS, to A.B., G.S. and A.D.)', the European Union ERA-Net for Research Programmes on Rare Diseases (NEUROLIPID, to G.S.), the VERUM Foundation (to A.B., A.D. and G.S.), the NIH (R01NS075764, 5R01NS072248 and U54NS065712 to S.Z.), the Fondation Lejeune (to C.O.) and benefited from the program 'Investissements d'Avenir' ANR-10-IAIHU-06 (to the ICM). M.C. was the recipient of a fellowship from the Fonds de la Recherche Scientifique (aspirant FNRS). S.M. was the recipient of a fellowship (SFRH/BD/87189/2012) from the FCT (Fundação para a Ciência e Tecnologia) supported by POPH/FSE funding.

Supplementary material

Supplementary material is available at *Brain* online.

References

Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin–rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet 2002; 30: 97–101.

Baumgartner MR, Hu CA, Almashanu S, Steel G, Obie C, Aral B, et al. Hyperammonemia with reduced ornithine, citrulline, arginine and proline: a new inborn error caused by a mutation in the gene encoding delta(1)-pyrroline-5-carboxylate synthase. Hum Mol Genet 2000; 9: 2853–8.

Baumgartner MR, Rabier D, Nassogne MC, Dufier JL, Padovani JP, Kamoun P, et al. Delta1-pyrroline-5-carboxylate synthase deficiency: neurodegeneration, cataracts and connective tissue manifestations combined with hyperammonaemia and reduced ornithine, citrulline, arginine and proline. Eur J Pediatr 2005; 164: 31–6.

Bicknell LS, Pitt J, Aftimos S, Ramadas R, Maw MA, Robertson SP. A missense mutation in ALDH18A1, encoding Delta1-pyrroline-5-carboxylate synthase (P5CS), causes an autosomal recessive neurocutaneous syndrome. Eur J Hum Genet 2008; 16: 1176–86.

Blackstone C. Cellular pathways of hereditary spastic paraplegia. Annu Rev Neurosci 2012; 35: 25–47.

DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 2011; 43: 491–8.

Dimopoulou A, Fischer B, Gardeitchik T, Schroter P, Kayserili H, Schlack C, et al. Genotype-phenotype spectrum of PYCR1-related autosomal recessive cutis laxa. Mol Genet Metab 2013; 110: 352–61.

Ebermann I, Phillips JB, Liebau MC, Koenekoop RK, Schermer B, Lopez I, et al. PDZD7 is a modifier of retinal disease and a contributor to digenic Usher syndrome. J Clin Invest 2010; 120: 1812–23.

Esteves T, Durr A, Mundwiller E, Loureiro JL, Boutry M, Gonzalez MA, et al. Loss of association of REEP2 with membranes leads to hereditary spastic paraplegia. Am J Hum Genet 2014; 94: 268–77.

Fichman Y, Gerdes SY, Kovacs H, Szabados L, Zilberstein A, Csonka LN. Evolution of proline biosynthesis: enzymology, bioinformatics, genetics, and transcriptional regulation. Biol Rev Camb Philos Soc 2014. doi: 10.1111/brv.12146.

Finsterer J, Loscher W, Quasthoff S, Wanschitz J, Auer-Grumbach M, Stevanin G. Hereditary spastic paraplegias with autosomal dominant, recessive, X-linked, or maternal trait of inheritance. J Neurol Sci 2012; 318: 1–18.

Fischer B, Callewaert B, Schroter P, Coucke PJ, Schlack C, Ott CE, et al. Severe congenital cutis laxa with cardiovascular manifestations due to homozygous deletions in ALDH18A1. Mol Genet Metab 2014; 112: 310–6.

Gardeitchik T, Mohamed M, Fischer B, Lammens M, Lefeber D, Lace B, et al. Clinical and biochemical features guiding the diagnostics in neurometabolic cutis laxa. Eur J Hum Genet 2014; 22: 888–95.

Gonzalez MA, Lebrigio RF, Van Booven D, Ulloa RH, Powell E, Speziani F, et al. GEnomes Management Application (GEM.app): a new software tool for large-scale collaborative genome analysis. Hum Mutat 2013; 34: 842–6.

Haberle J, Boddaert N, Burlina A, Chakrapani A, Dixon M, Huemer M, et al. Suggested guidelines for the diagnosis and management of urea cycle disorders. Orphanet J Rare Dis 2012; 7: 32.

Handley MT, Megarbane A, Meynert AM, Brown S, Freyer E, Taylor MS, et al. Loss of ALDH18A1 function is associated with a cellular lipid droplet phenotype suggesting a link between autosomal recessive cutis laxa type 3A and Warburg Micro syndrome. Mol Genet Genomic Med 2014; 2: 319–25.

Hu CA, Bart Williams D, Zhaorigetu S, Khalil S, Wan G, Valle D. Functional genomics and SNP analysis of human genes encoding proline metabolic enzymes. Amino Acids 2008a; 35: 655–64.

Hu CA, Khalil S, Zhaorigetu S, Liu Z, Tyler M, Wan G, et al. Human Delta1-pyrroline-5-carboxylate synthase: function and regulation. Amino Acids 2008b; 35: 665–72.

Hu CA, Lin WW, Obie C, Valle D. Molecular enzymology of mammalian Delta1-pyrroline-5-carboxylate synthase. Alternative splice donor utilization generates isoforms with different sensitivity to ornithine inhibition. J Biol Chem 1999; 274: 6754–62.

Kemp S, Berger J, Aubourg P. X-linked adrenoleukodystrophy: clinical, metabolic, genetic and pathophysiological aspects. Biochim Biophys Acta 2012; 1822: 1465–74.

Koch MC, Steinmeyer K, Lorenz C, Ricker K, Wolf F, Otto M, et al. The skeletal muscle chloride channel in dominant and recessive human myotonia. Science 1992; 257: 797–800.

Martinelli D, Haberle J, Rubio V, Giunta C, Hausser I, Carrozzo R, et al. Understanding pyrroline-5-carboxylate synthetase deficiency: clinical, molecular, functional, and expression studies, structure-based analysis, and novel therapy with arginine. J Inherit Metab Dis 2012; 35: 761–76.

Novarino G, Fenstermaker AG, Zaki MS, Hofree M, Silhavy JL, Heiberg AD, et al. Exome sequencing links corticospinal motor neuron disease to common neurodegenerative disorders. Science 2014; 343: 506–11.

Rabier D, Diry C, Rotig A, Rustin P, Heron B, Bardet J, et al. Persistent hypocitrullinaemia as a marker for mtDNA NARP T 8993 G mutation?. J Inherit Metab Dis 1998; 21: 216–9.

Reversade B, Escande-Beillard N, Dimopoulou A, Fischer B, Chng SC, Li Y, et al. Mutations in PYCR1 cause cutis laxa with progeroid features. Nat Genet 2009; 41: 1016–21.

Schule R, Siddique T, Deng HX, Yang Y, Donkervoort S, Hansson M, et al. Marked accumulation of 27-hydroxycholesterol in SPG5 patients with hereditary spastic paresis. J Lipid Res 2010; 51: 819–23.

Sedel F, Fontaine B, Saudubray JM, Lyon-Caen O. Hereditary spastic paraparesis in adults associated with inborn errors of metabolism: a diagnostic approach. J Inherit Metab Dis 2007; 30: 855–64.

Skidmore DL, Chitayat D, Morgan T, Hinek A, Fischer B, Dimopoulou A, et al. Further expansion of the phenotypic spectrum associated with mutations in ALDH18A1, encoding Delta(1)-pyrroline-5-carboxylate synthase (P5CS). Am J Med Genet A 2011; 155A: 1848–56.

Synofzik M, Gonzalez MA, Lourenco CM, Coutelier M, Haack TB, Rebelo A, et al. PNPLA6 mutations cause Boucher-Neuhauser and Gordon Holmes syndromes as part of a broad neurodegenerative spectrum. Brain 2014; 137 (Pt 1): 69–77.

Tesson C, Koht J, Stevanin G. Delving into the complexity of hereditary spastic paraplegias: how unexpected phenotypes and inheritance modes are revolutionizing their nosology. Hum Genet 2015; 134: 511–38.

Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 2010; 38: e164.

Wolthuis DF, van Asbeck E, Mohamed M, Gardeitchik T, Lim-Melia ER, Wevers RA, et al. Cutis laxa, fat pads and retinopathy due to ALDH18A1 mutation and review of the literature. Eur J Paediatr Neurol 2014; 18: 511–5.

Yasuda T, Kaji Y, Agatsuma T, Niki T, Arisawa M, Shuto S, et al. DJ-1 cooperates with PYCR1 in cell protection against oxidative stress. Biochem Biophys Res Commun 2013; 436: 289–94.

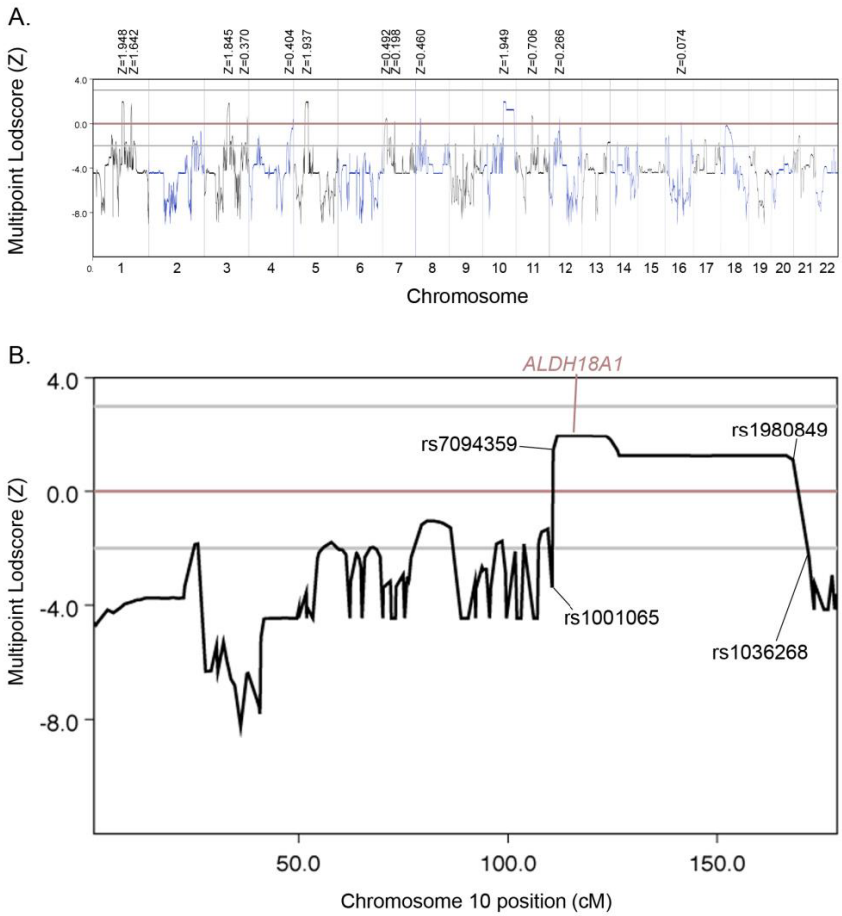
Zampatti S, Castori M, Fischer B, Ferrari P, Garavelli L, Dionisi-Vici C, et al. De Barsy syndrome: a genetically heterogeneous autosomal recessive cutis laxa syndrome related to P5CS and PYCR1 dysfunction. Am J Med Genet A 2012; 158A: 927–31.

Supplementary Figure 1 Whole genome linkage analysis in family FSP410

(A) Multipoint linkage analysis on all the autosomes in family FSP410. Whole genome linkage analysis was performed using Illumina LINKAGE_12 microarrays (6090 SNP markers). Genotypes were determined using Beadstudio (Illumina) and analyzed with MERLIN 1.0 (Abecasis *et al.*, 2002) assuming an autosomal dominant transmission under a 0.80 penetrance model with equal allele frequencies, similar recombination fractions between males and females, and a disease frequency of 0.00001. Genotypes were available for 13 individuals, including affected subjects 6, 8, 12, 13, 16, 25, 29 and their relatives, 5, 7, 9, 14, 19, 24. Whole genome linkage analysis could not lead to the identification of a unique well-defined associated locus due to pedigree structure and size limitations. Instead, various putatively linked or unexcluded (LOD scores above -2) loci were found in the family, including five putatively linked loci with multipoint LOD scores reaching the maximal expected values of this pedigree (from +1.64 to +1.95) and various uninformative regions.

(B) Close-up view of the linkage analysis results for chromosome 10. A significant multipoint LOD score (Zmax = +1.949) was found between markers rs1001065 and rs1036268, in a region encompassing the *ALDH18A1* gene.

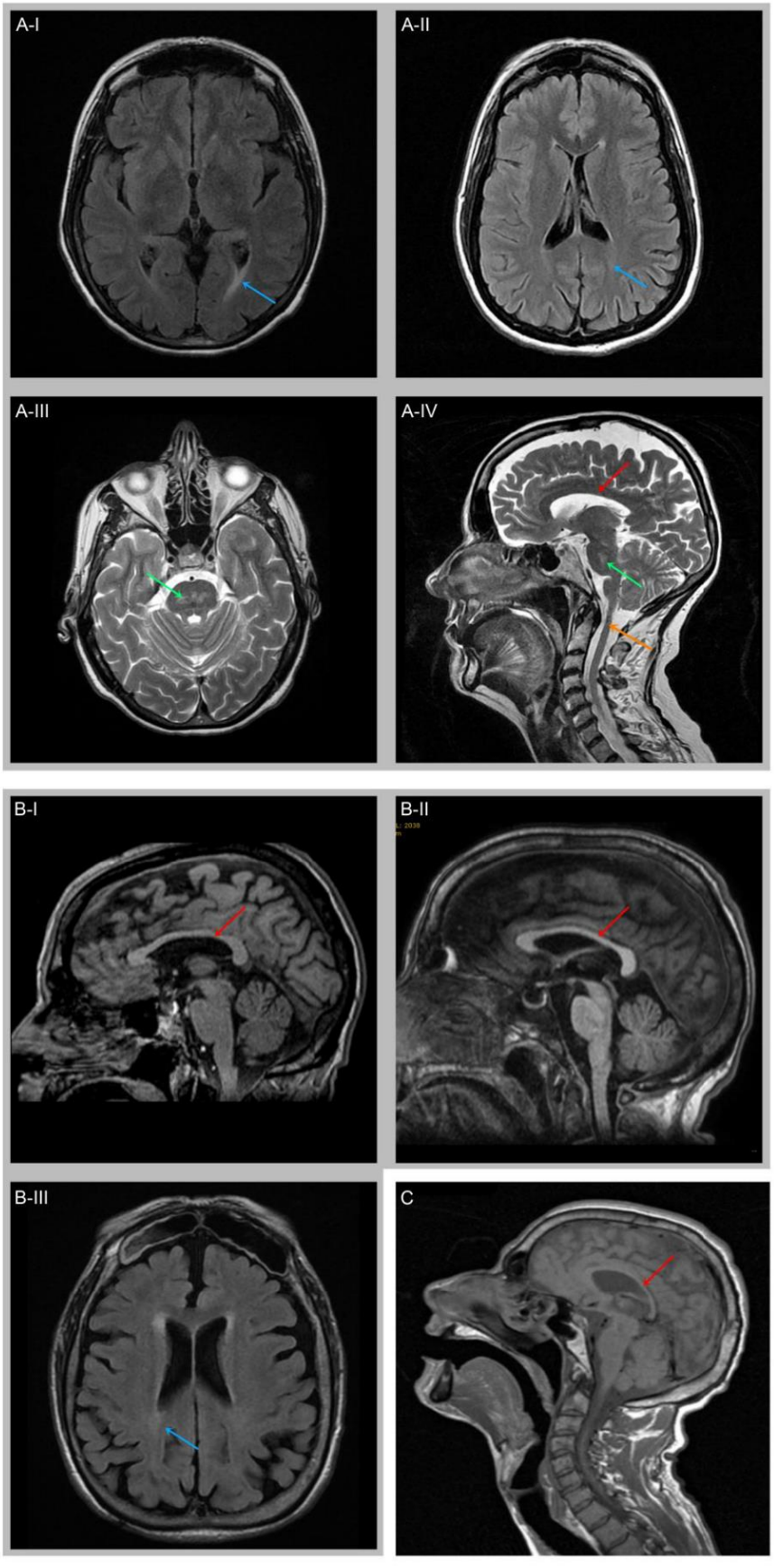
Reference: Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet. 2002;30(1):97-101.



Supplementary Figure 2 Brain magnetic resonance imaging

MRI was performed for both patients from family FSP429 (upper grey background), FSP429-6 (panel A-I: FLAIR axial section, III: T2 axial section, IV: T2 sagittal section) and FSP429-21 (panel A-II: FLAIR axial section); in two patients from family FSP410 (lower grey background), FSP410-32 (panel B-I: T2 sagittal section) and FSP410-13 (panel B-II: T2 sagittal section, III: FLAIR axial section); and in patient SR45-II:6 (panel C: T1 sagittal section).

Diverse anomalies are described. A thin corpus callosum (TCC), previously reported in carriers of *ALDH18A1* biallelic mutations, was observed (red arrows) in patient SR45-II:6 (C); moderate TCC is also found in patient FSP410-32 (B-I) and FSP410-13 (B-II), while global cerebral atrophy including the corpus callosum is present in patient FSP429-6 (A-IV). Moderate spinal cord atrophy (orange arrow) is described in patient FSP429-6 (A-IV) as well as patient FSP429-21 and sporadic patient GSHSP44 (not shown). Mild white matter anomalies (blue arrows) are reported in patients FSP429-6 (A-I), FSP429-21 (A-II) and FSP410-13 (B-III). T2 hyperintensity (green arrows), possibly of vascular origin, is noticed in the pons of patient FSP429-6 (A-III, IV).

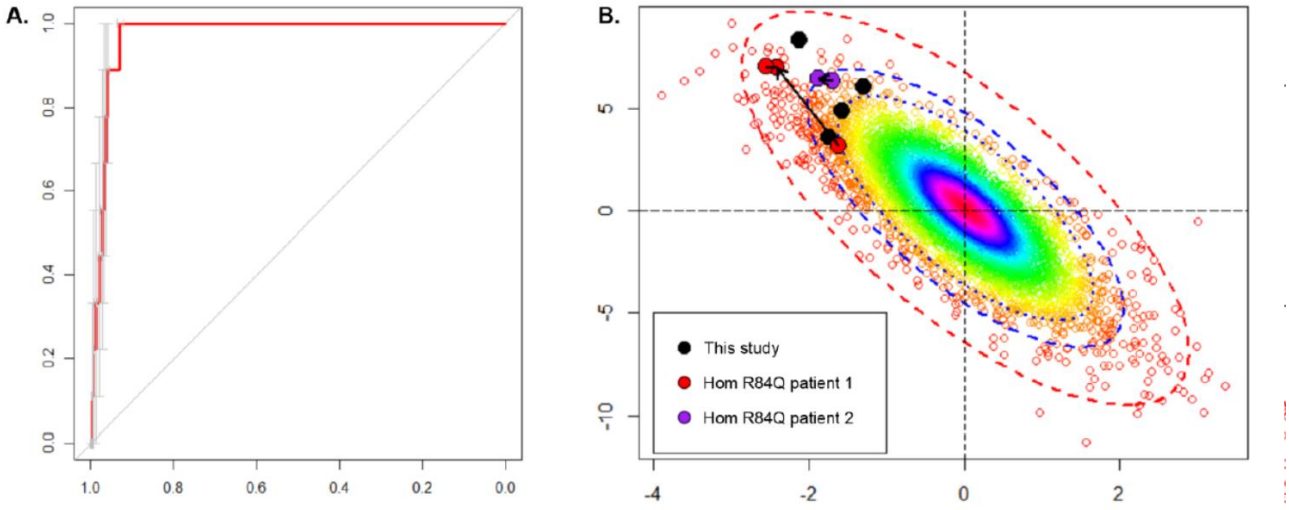

A-I
A-II
A-III
A-IV
B-I
B-II
B-III
C

Supplementary Figure 3 Biochemical analyses as predictor of the genetic status

(A) Prediction of mutation status for *ALDH18A1* based on the mean of age-normalized values for plasma proline, arginine, ornithine and citrulline. The patients from this study and the first family with reported mutations in *ALDH18A1* (Baumgartner et al, 2005) are compared to the full cohort of adult patients (N=5023). The resulting area under the receiver operating characteristic (ROC) curve is 97%, with 95% confidence interval: 96-99%.

(B) We plotted two linear combinations of plasma amino acid markers to further compare the full cohort of adult patients of Necker hospital (N=5023, red circles) *vs* the *ALDH18A1*-mutated patients (see figure legend). X-axis, the mean of age-normalized values for plasma proline, arginine, ornithine and citrulline. Y-axis, plasma glutamine minus the mean of age-normalized values for plasma proline, ornithine and citrulline. The ellipses indicate alpha error cut-offs (p-values) of 0.05, 0.01 and 0.0001 based on F-distribution (dotted blue, dashed blue and dashed red lines). The patients with *ALDH18A1* mutations are all located between $p<0.05$ and $p=0.0001$. The distribution suggests that the Y-axis improves discrimination relative to the X-axis alone.

Reference: Baumgartner MR, Rabier D, Nassogne MC, Dufier JL, Padovani JP, Kamoun P, et al. Delta1-pyrroline-5-carboxylate synthase deficiency: neurodegeneration, cataracts and connective tissue manifestations combined with hyperammonaemia and reduced ornithine, citrulline, arginine and proline. Eur J Pediatr. 2005;164(1):31-6



Supplementary Table 1 Primers used for *ALDH18A1* panel sequencing

Exon	Forward primer	Reverse primer
2	AAGACTCGGCAGCATCTCCAGGTAAGGGGAAGTAGAGGGAAA	GCGATCGTCACTGTTCTCCATCAAACGTCCCACATACTCA
3	AAGACTCGGCAGCATCTCCATGGTTAATACTGAGCATGTATATTTTT	GCGATCGTCACTGTTCTCCAAAAAACCCAATTTTAGTACAGCA
4	AAGACTCGGCAGCATCTCCACCTACCTCCATCCCTTGTGA	GCGATCGTCACTGTTCTCCAGCTCGTGCCTCTAAGACTGG
5	AAGACTCGGCAGCATCTCCATTCCCTCATTCCCTACCACA	GCGATCGTCACTGTTCTCCAGTCCCACACCCATTTCAAAG
6	AAGACTCGGCAGCATCTCCATTCTCCCTCATGCATTTTCC	GCGATCGTCACTGTTCTCCAATGCAGCATAGCATGGTGTT
7	AAGACTCGGCAGCATCTCCATTGGCTGTGTTCTTCCATGA	GCGATCGTCACTGTTCTCCAGCATTTAAAACAAAGGGCAGA
8	AAGACTCGGCAGCATCTCCACAGGTGGTAGTGCAGTGGAA	GCGATCGTCACTGTTCTCCACATTAACCCCCAAATAATTACAAAA
9	AAGACTCGGCAGCATCTCCATGCAAGTTTAAAATGTGGGTGA	GCGATCGTCACTGTTCTCCAGCGTTCAAAACAATGCCATA
10	AAGACTCGGCAGCATCTCCAGCTTGTGCATGTGCTGGTAG	GCGATCGTCACTGTTCTCCATCAGCAGGTTTGACTCCTGT
11	AAGACTCGGCAGCATCTCCAAGGTTGGCAGTGAGAGAGGA	GCGATCGTCACTGTTCTCCAAAATAATGCACACCCCTCCA
12	AAGACTCGGCAGCATCTCCAATCGGTTGGTAGCTGTAGCC	GCGATCGTCACTGTTCTCCACTTCCTCCAATATCCACACCA
13	AAGACTCGGCAGCATCTCCATGTCTCCTTTGATCCACTGC	GCGATCGTCACTGTTCTCCAGCTCTGCAGAATGGCCTAAA
14	AAGACTCGGCAGCATCTCCATCTGTAAAAGGGAAGCTGCTG	GCGATCGTCACTGTTCTCCATGAAAGAGAAGACCCCATCC
15	AAGACTCGGCAGCATCTCCAGGATCTGTCACATGTATGTGGATT	GCGATCGTCACTGTTCTCCAGCTGTGCCTGGTCTAATTCC
16	AAGACTCGGCAGCATCTCCATTGGGCTGTGGTTTTACAGG	GCGATCGTCACTGTTCTCCAGGATCAGAAAGCAGCAAAGG
17	AAGACTCGGCAGCATCTCCAAGTCAATGGCTGGAAACAGG	GCGATCGTCACTGTTCTCCAACCCCTAGACCACACAGTGC
18	AAGACTCGGCAGCATCTCCAAGAGGCTGGGAAAAGCTACA	GCGATCGTCACTGTTCTCCATTCCAACAGGCAGACCCTAC

The specific sequence is indicated in red. The adaptor sequence that allows for subsequent sequencing is indicated in blac

4.2.2 *AIMP1* as a causative gene for HSPs

Title: *AIMP1*, a gene implicated in a range of phenotypes from intellectual disability to hereditary spastic paraplegia

Short running title: *AIMP1*, a new hereditary spastic paraplegia gene

Authors: Sara Morais, MSc (1-8), Christelle Tesson, PhD (4-8), Paula Coutinho, MD (1,2), Eva Brandão, MD (10), Paula Ribeiro, MD (10), José Leal Loureiro, MD, PhD (1,2,10), Jorge Sequeiros, MD, PhD (1-3), Alexis Brice, MD, PhD (4-7,9), Giovanni Stevanin, PhD (4-9), Isabel Alonso, PhD (1-3)

Affiliations:

1- UnIGENe, Instituto de Biologia Molecular e Celular (IBMC), Universidade do Porto, Portugal
2- Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Portugal
3- Instituto de Ciências Biomédicas de Abel Salazar (ICBAS), Universidade do Porto, Portugal
4- INSERM, U 1127, F-75013, Paris, France
5- CNRS, UMR 7225, F-75013, Paris, France
6- Sorbonne Universités, UPMC Univ Paris 06, UMRS_1127, F-75013, Paris, France
7- Institut du Cerveau et de la Moelle épinière, ICM, F-75013, Paris, France
8- Ecole Pratique des Hautes Etudes, EPHE, PSL Research University, F-75013, Paris, France
9- APHP, Hôpital de la Pitié-Salpêtrière, Département de Génétique, F-75013, Paris, France
10- Serviço de Neurologia, Centro Hospitalar de Entre o Douro e Vouga, Santa Maria da Feira, Portugal

Correspondence: Dr I Alonso, UnIGENe, Instituto de Biologia Molecular e Celular (IBMC), Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Rua Alfredo Allen, 208, Porto 4200-135, Portugal. Tel: +351 220 408 800; E-mail: ialonso@ibmc.up.pt

Abstract

Purpose: Hereditary spastic paraplegias (HSP) are a group of neurodegenerative disorders clinically and genetically heterogeneous, with lower limb spasticity and weakness as the core phenotype. Up to now, more than 70 genes have been associated with the disease, but many families still remain without molecular diagnosis.

Methods: Linkage analysis and exome sequencing was performed to identify the disease-causing gene in a large HSP family with recessive inheritance. A minigene construct was created in order to confirm a predicted splicing effect of the identified variant.

Results: Multipoint linkage analysis showed a maximum LOD score of 4.3 in chromosome 4, which, after exome sequencing, led to the identification of a disease-causing *AIMP1* variant, c.295G>A, in homozygosity. The minigene strategy showed that the apparent missense variant, p.(Val99Met), causes the loss of exon 3 donor splice-site, resulting in the translation of part of this intron 3 and the introduction of a premature stop codon.

Conclusions: The identification of *AIMP1* disease-causing variants in a HSP family adds to the large number of HSP-related genes identified and shows that *AIMP1* disruption seems to be responsible for a range of disorders, from isolated intellectual disability to complex neurodegenerative diseases.

Key Words: Minigene strategy, splice-site loss, next generation sequencing, clinical heterogeneity, neurodegenerative disease

Introduction

Hereditary spastic paraplegias (HSP) are a group of rare neurodegenerative disorders characterized by spasticity and weakness of the lower limbs, and which may be complicated by other neurological or extra-neurological symptoms like cerebellar ataxia, cognitive impairment and neuropathy (Fink, 2013). More than 70 genes have already been associated with the disease with new ones constantly being added, though numerous families remain without molecular diagnosis (Tesson et al., 2015).

A population-based survey conducted in Portugal, from 1994 to 2004, allowed the identification of a large cohort of families with HSP (Coutinho et al., 2013; Silva et al., 1997). A large genetic screening was performed in these families (Morais et al., 2017), but in almost half of them (49%) the causing gene remains to be identified. A large family with consanguinity was selected for linkage analysis and exome sequencing. This led to the identification of a homozygous variant in *AIMP1* (aminoacyl tRNA synthetase complex interacting multifunctional protein 1) as responsible for the phenotype in this family.

AIMP1/p43 function as a non-catalytic component of the multi-tRNA synthetase complex. Mouse studies showed that it is expressed in neurons of the central nervous system and required for the maintenance of axonal integrity, thus being predictable to find its dysfunction associated with neurodegenerative diseases (Zhu et al., 2009). *AIMP1* disease-causing variants were previously associated with Pelizaeus-Merzbacher-like disease (PMLD) (Feinstein et al., 2010), an autosomal recessive leukodystrophy that is a differential diagnosis for HSPs (Tesson et al., 2015). Debate about the first published case and reports of new families extended its clinical spectrum from neurodegenerative disease to intellectual disability (Armstrong et al., 2014; Biancheri et al., 2011; Boespflug-Tanguy et al., 2011; Iqbal et al., 2016). Here, we report the first HSP family with an *AIMP1* disease-causing variant, further expanding the phenotype related to this gene.

Materials and Methods

Patient recruitment and clinical evaluation

The family selected for this study (SR9, Figure 1) was recruited during the population-based survey in Portugal (Coutinho et al., 2013; Silva et al., 1997), where all families were evaluated by the same team of neurologists using homogeneous inclusion criteria. Written informed consent for genetic study was obtained and DNA stored. Being a large family with two consanguineous couples and DNA available from several individuals, previously excluded for commonly mutated genes in recessive HSP (*SPG11*, *ZFYVE26* and *CYP7B1*), it was selected for a genome-wide screen and linkage analysis. Clinical features of the three patients (2 brothers and a cousin) are summarized in supplementary Table 1.

Whole genome linkage analysis and exome sequencing

A genome-wide linkage scan was performed, using 300 microsatellite markers spaced 15cM apart on autosomes (as described in (Stevanin et al., 2007b)). Multipoint linkage analyses, using Allegro 1.2c (Gudbjartsson et al., 2000), resulted in a peak LOD score of 4.3 on chromosome 4. Exome sequencing was performed at the ICM genetic platform (IgenSeq), using the Nextera Rapid Capture Exome kit (Illumina, San Diego, USA) and the NextSeq500 sequencer (Illumina, San Diego, USA), on patients VI:3 and VI:6. For data analysis, a custom workflow was applied, using FastQC for quality control, BWA-0.7.8 for alignment with the hg19 reference genome, GATK for coverage analysis and variant calling, and ANNOVAR for annotation. Exome data was then filtered, by focusing on homozygous variants present in the defined linkage region. To be further considered, variants needed to cause a change at the protein level and, if known in databases, its minor allele frequency (MAF) had to be below 1%. All filtered variants were further analysed using Alamut v.2.4 (Interactive Biosoftware, La Rochelle, France) for pathogenicity prediction and its frequency checked in the Exome Aggregation Consortium (ExAC) database **(**http://exac.broadinstitute.org/**).**

AIMP1 minigenes

The *AIMP1* minigene was obtained by cloning exons 2 to 4 with intronic regions (amplified from the patient sample) into the pCMVdi vector, kindly provided by Dr Alexandra Moreira (Glória et al., 2014), by Gibson Assembly. To generate the wild-type construct, we performed site-directed mutagenesis, using the QuikChange Site-Directed Mutagenesis Kit (Agilent Technologies, Santa Clara, USA), according to the manufacturer's protocol, with the appropriate primer pairs designed with the QuikChange Primer Design Program (Agilent Technologies, Santa Clara, USA). HEK293T cells were transfected with the created constructs using jetPrime (Polyplus, Berkeley, USA) transfection reagent. RNA was extracted 48h after transfection, by a double method with NZYol reagent (Nzytech, Lisbon, Portugal) and RNeasy Mini Kit (QIAGEN, Hilden, Germany). Synthesis of cDNA by reverse transcription was performed with SuperScript III First-Strand Synthesis System (Invitrogen, Carlsbad, USA), according to manufacturer's instructions. PCR amplification of the resulting cDNA from exons 2 to 4 and Sanger sequencing were performed as described below.

Sanger sequencing

Sanger sequencing was performed to confirm the variants presence and their segregation in the family. Additionally, we sequenced the seven *AIMP1* exons in 67 Portuguese HSP families that remained without diagnosis (Morais et al., 2017). Primers (available upon request) were designed with the Primer3Plus software and PCR amplification was done using the HotStart

Master Mix (Qiagen, Hilden, Germany) followed by Exo/SAP Go (Grisp, Porto, Portugal) purification and sequencing, using BigDye Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, USA), in an ABI-PRISM 3130 XL automatic sequencer (Applied Biosystems, Foster City, USA).

Results

Linkage analysis of a large HSP family using microsatellite markers revealed a region of linkage, on chromosome 4, with a maximum LOD score of 4.3 (Figure 1B), ranging from D4S1591 to D4S2975. This spans approximately 23.2 Mb and includes 87 coding genes. Exome analysis (focused on the region of interest) showed 3 missense homozygous variants in *AIMP1*, *ALPK1* and *KIAA1109* with minor allele frequency (MAF) below 1% and segregating with the disease. Bioinformatics analysis excluded the variants *KIAA1109* and *ALPK1*, although *ALPK1* is a gene associated with motor coordination deficits in mutant mice (Chen and Xu, 2011), due to their MAF (>0.6%) and presence in the ExAC database in ≥5 individuals in the homozygous state. The *AIMP1* variant was predicted to be pathogenic by Mutation Taster and Polyphen-2. Interestingly, it was also predicted to affect splicing by NNSplice and GeneSplicer; and it was found in the ExAC database only in 2 alleles (no homozygotes). As the only two reported variants of *AIMP1* causing neurodegeneration were a frameshift and a nonsense variant leading to truncated forms of the protein, we decide to further investigate the potential splicing effect of this variant. To accomplish this, we created a minigene that might theoretically mimic *in vivo* conditions, by cloning exons 2 to 4 with intronic regions 2 and 3 in between, into the pCMVdi vector (Glória et al., 2014); the wild-type construct was obtained by site-directed mutagenesis (Figure 1D-G). After transfection and cDNA sequencing, we confirmed that variant c.295G>A, which changes the last base of exon 3, causes the loss of the donor splice-site, leading to the inclusion of 74 intronic bases into exon 3, where an alternative splice site is found. This leads to a change of valine (at position 99) into serine, and not a methionine as predicted, and the insertion of several bases and of a premature stop codon 2 amino acids later. The 67 additional Portuguese HSP families, previously excluded for common recessive HSP, were screened by Sanger sequencing of all exons and intronic flanking regions, but no additional families with *AIMP1* disease-causing variants were found.

Discussion

Exome analysis found a single homozygous variant in *AIMP1* as the only possible causative variant in the linkage region identified. This gene was associated with Pelizaeus-Merzbacher-like Disease (PMLD) (Feinstein et al., 2010), which shares several causative genes with HSP, such as *GJC2*, *HSPD1*, *SLC16A2* and *PLP1*. Although the clinical diagnosis of the first reported case was a matter of debate (Biancheri et al., 2011; Boespflug-Tanguy et al., 2011), evidence still indicates *AIMP1* as the causal gene in a family that also presented symptoms of spastic paraplegia.

Clinically, all three affected members (Figure 1A) of our family SR9 presented with severe spastic paraplegia in lower limbs and moderate in upper limbs, severe gait impairment, dysphagia and dysarthria. They had delayed motor development (2/3) and cognitive impairment. Pes cavus and distal muscle wasting were present in all patients (Supplementary Table 1), suggesting involvement of the second motor neuron. In addition to an Israeli family with a homozygous frameshift variant due to a deletion, associated with PMLD (Feinstein et al., 2010), it was also reported a homozygous nonsense variant in a Filipino patient with a cortical neurodegenerative disease and a rather severe phenotype, leading to premature dead (Armstrong et al., 2014). More recently, two families (from Pakistan and Iran) were reported with a homozygous missense variant causing intellectual disability, without neurodegenerative disease (Iqbal et al., 2016). Comparing the previously reported families with *AIMP1* variants (Table 1), all share early onset and intellectual disability. Also, it seems that truncating variants are associated with a more complex and severe phenotype. This is the case of our apparently missense variant, which in fact is a splicing variant responsible for the insertion of a premature stop codon that can either lead to the target of mRNA to nonsense mediated decay or result in a truncated protein.

AIMP1/p43 is a cofactor of the mammalian aminoacyl-tRNA synthetase complex, which seems to be essential for neurofilaments assembly and axon maintenance, by keeping the normal phosphorylation levels. Abnormal accumulation of hyperphosphorylated neurofilaments is a pathological hallmark of several neurodegenerative disorders, as Parkinson disease and amyotrophic lateral sclerosis. AIMP1 protein was shown to be predominantly expressed in central nervous system neurons and to interact with neurofilament light subunit *in vivo*. Mice lacking this protein showed disorganization of the axonal neurofilament network and exhibit motor symptoms and axonal degeneration of motor neurons (Zhu et al., 2009), what seems to be in accordance with the motor phenotype also present in patients.

This study identifies *AIMP1* as a new HSP disease-causing gene that seems to be involved in a large range of phenotypes, with a possible genotype-phenotype correlation in which truncating variants result in neurodegenerative disorders and missense variants being the cause of isolated intellectual disability.

Acknowledgments

We would like to thank the participating patients and families, and their clinicians; Emeline Mundwiller and Delphine Bouteiller for their help in the exome sequencing; and Justine Guégan for the help with the exome data. This study was funded in part by FCT (Fundação para a Ciência e Tecnologia, grant FCT-ANR/BEX-GMG/0008/2013, to IA) and the French National Agency for Research (SPATAX-QUEST, to GS), the European Union (Omics call of the 7th framework program, NEUROMICS, to AB; the E-rare program, NEUROLIPID and PREPARE, to GS), the VERUM Foundation (to AB and GS); and benefited from the program “Investissements d’Avenir” ANR-10-IAIHU-06 (to the ICM). SM is the recipient of a fellowship (SFRH/BD/87189/2012) from FCT, supported by POPH/FSE funding.

Conflicts of Interest

The authors report no conflicts of interest.

Table 1 Major clinical features of the families published with *AIMP1* disease causing variants in homozygosity

Reference	(Feinstein et al., 2010)	(Armstrong et al., 2014)	(Iqbal et al., 2016)		**This study (SR9)**
Variant*	c.364_365delCA	c.187C>T	c.967G>A	c.599T>G	c.295G>A
Protein effect	p.(Gln122Valfs*30)	p.(Gln63*)	p.(Gly323Arg)	p.(Val200Gly)	p.(Val99Met) -> p.(Val99Serfs*2)
Age-at-onset	<1	<1	NA	~2	1
Spastic paraparesis	Yes	NA	-	-	Yes
Intellectual disability/ Development delay	Yes	Yes	Yes	Yes	Yes
Growth retardation/ Microcephaly	Yes/Yes	Yes/Yes	Yes (1/5)/No	No/No	No
Hypotonia	Yes	Yes	Yes (neck)	No	No
Pyramidal signs	Yes	-	-	-	Yes
Other neurological findings	Seizures	Seizures	-	-	Distal muscle wasting
Ophthalmological findings	Nystagmus	No reaction to light	-	Myopia (1/2)	-
MRI	Myelin deficiency Corpus callosum atrophy	Myelin deficiency Cerebellar atrophy	Normal	NA	NA

***** All described variants were reported according to transcript NM_001142415.1 that are here identified according to transcript NM_001142416

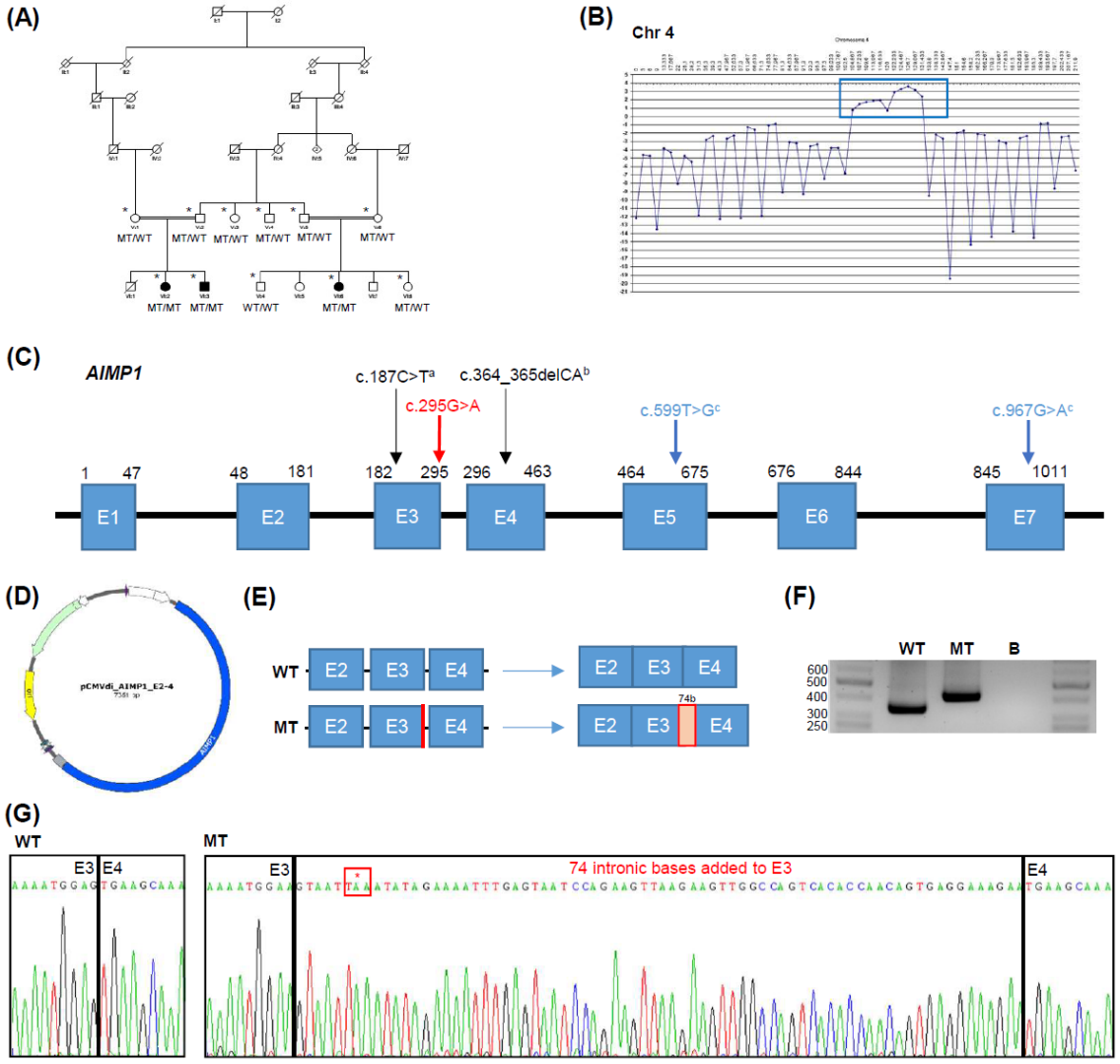


Figure 1 From pedigree to *AIMP1* identification and variant analysis. (A) Family pedigree with segregation analysis. An asterisk indicates family members with DNA available. (B) Multipoint linkage data from chromosome 4, with a blue square indicating the linkage region identified, with a maximum LOD score of 4.3. (C) Scheme of the *AIMP1* gene with previously reported variants on the NM_001142416 transcript. Truncating variants are marked in black, missense in blue, and the variant identified in our family in red. [a-] (Armstrong et al., 2014), [b-] (Feinstein et al., 2010), [c-] (Iqbal et al., 2016). (D) Map of the plasmid created with the pCMV promoter and the *AIMP1* insert, to obtain a minigene. (E) Schematics of the exons and introns cloned into the plasmid, and results following HEK293T transfection and mRNA extraction. (F) PCR amplification of cDNA produced from RNA isolated from HEK293T cells, 48h after transfection with WT and MT plasmids. (G) Electropherogram obtained after cDNA amplification, with the exons junction marked and the position of the premature stop codon boxed in the 74 intronic bases not removed during splicing.

Supplementary information

Table S1 Clinical features of patients from family SR9 with homozygous c.295G>A variant in *AIMP1*

Family code	**SR9**		
Patient (gender)	VI:2 (F)	VI:3 (M)	VI:6 (F)
Phenotype	Complex	Complex	Complex
Age at onset (years)	1	1	1,5
Age at examination (years)	39	37	30
Disease duration (years)	38	36	29
Spasticity LL (0-4)	4	4	4
Spasticity UL (0-4)	1	1	1
Weakness LL (0-5)	2	2	2
Weakness UL (0-5)	4	4	4
Babinski sign	Yes	No	Yes
Sphincter disturbances	NA	NA	No
Dysarthria and dysphagia	Yes	Yes	Yes
Cognitive impairment	Yes	Yes	Yes
Other signs	Pes cavus, distal muscle wasting	Pes cavus, distal muscle wasting	Pes cavus, distal muscle wasting
MRI	NA	NA	NA

M, male; F, female; LL, lower limbs; UL, upper limbs; NA, not available; MRI, magnetic resonance imaging.

For spasticity a scale of 0 to 4, according to the Modified Ashworth Scale, was used. Weakness was graded from 0 to 5, using the Medical Research Council Scale for Muscle Strength.

4.2.3 *AKAP6* as a candidate for HSPs

A-Kinase Anchoring Protein 6 (AKAP6) disease-causing variants identified as a new cause for hereditary spastic paraplegia

In Portugal, a large cohort of 193 families with HSP was identified in a population-based survey (Coutinho et al., 2013; Silva et al., 1997). One of these families was the large SR2 family where the SPG32 locus was identified by linkage analysis (Stevanin et al., 2007b). Exome sequencing of this family led to the discovery of a homozygous variant in the *AKAP6* gene considered as disease-causing.

AKAP6 encodes the muscle A-kinase anchoring protein (mAKAP), a scaffolding protein found predominantly in the heart, skeletal muscle and brain. Previously known as AKAP100 due to the size of the original protein detected (McCartney et al., 1995), was then renamed mAKAP due to its presence in the striated muscle (Kapiloff et al., 1999). As an AKAP, it is mostly recognized for its interaction with the cAMP-dependent protein kinase (PKA), however as a scaffold for a multiple-enzyme complex and in the center of a large signalosome, it interacts also with a large number of signaling molecules. This include cAMP-specific phosphodiesterase type 4D3 (PDE4D3), type 2 ryanodine receptor (RyR2), adenylyl cyclase type 5 (AC5), protein phosphatases PP2A and calcineurin Aβ, phospholipase Cε, 3-phosphoinositide-dependent protein kinase-1 (PDK1) and nesprin-1α, among others. It is the binding with nesprin-1α, a protein inserted into the nuclear envelope due to a transmembrane domain, that results in a perinuclear localization of mAKAP (Pare et al., 2005). In mouse, mAKAP has two alternative-spliced isoforms that seem to be tissue-specific: mAKAPα, coding a 255kDa protein from the full-length mRNA sequence mostly found in the brain and mAKAPβ, preferentially expressed in the heart and skeletal muscle coding a 230kDa protein translated from an internal start site at position 245 (Michel et al., 2005; Passariello et al., 2015).

This study is the first report of a disease-causing variant in *AKAP6* and the first association of this gene with a neurodegenerative disease. Here we provide preliminary evidence gathered for a neuronal effect of this gene and of a pathogenic effect of this variant in the mAKAPα protein.

Results

Exome sequencing

The SPG32 locus, mapped to chromosome 14q12-q21, was identified as the largest homozygous region in three patients of a consanguineous Portuguese family with HSP. In

order to identify the gene and variant responsible for the disease, we studied two (of the three) patients by whole exome sequencing. The two exomes had a mean coverage of 110 and 136 reads with 93% of the bases covered above 20X. Exome analysis identified only one homozygous missense variant in the identified loci with a minor allele frequency (MAF) inferior to 1% and segregating with the disease in the family. The variant identified, c.464G>A, is present in the *AKAP6* gene, which leads to the alteration of the arginine 155 to a glutamine (Figure 1) and is predicted to be pathogenic by SIFT, Mutation Taster, Polyphen-2 and UMD-Predictor software. This variant is present in a conserved residue and, according to HOPE, wild type (WT) and mutant (MT) residues differ in size and charge. The altered residue is smaller, which might cause loss of potential interactions and also, as a neutral residue, causes the loss of the positive charge of the WT residue, which can be important for the interaction with other molecules or residues.

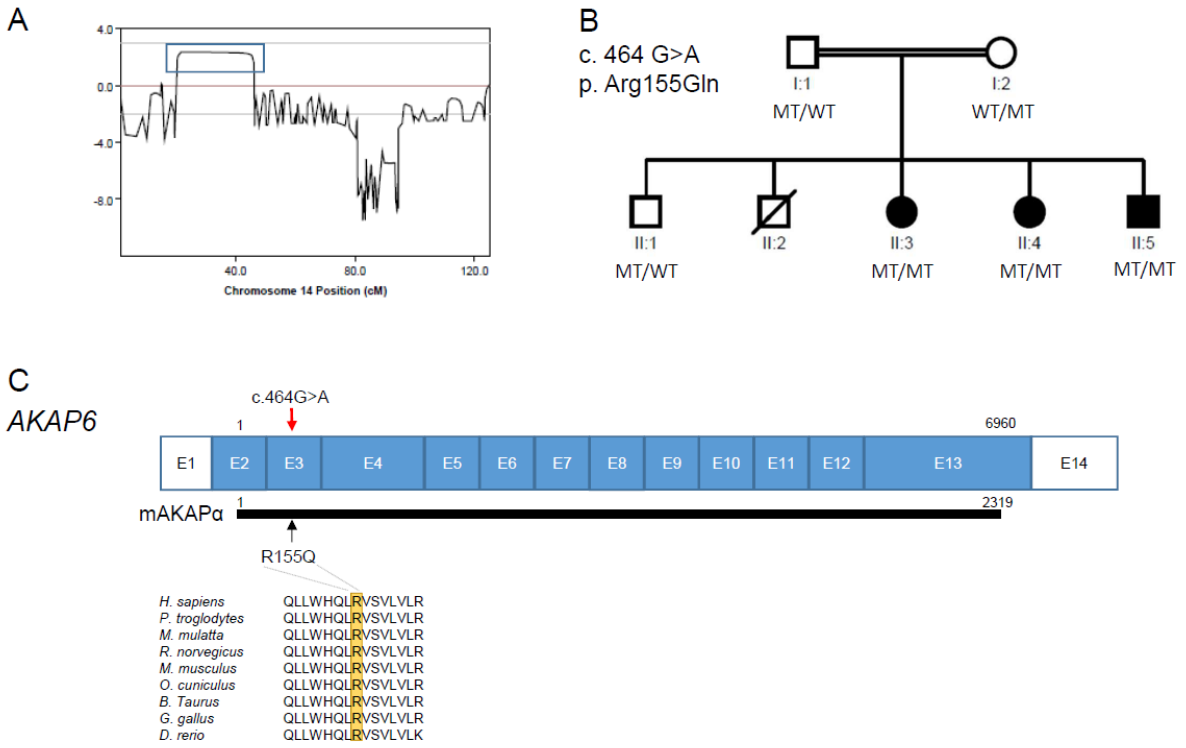


Figure 1. From linkage analysis to *AKAP6* variant identification. (A) Multipoint linkage data from chromosome 14 with a blue square indicating the identified linkage region with a significant multipoint LOD score peak of 3.74. (B) Family pedigree with segregation analysis. A black circle and square indicates affected individuals with HSP. (C) Schematics of *AKAP6* gene with the identified variant marked (NM_004274) and its protein conservation shown. Only the mAKAPα isoform is depicted here, as there is no correspondence between the mouse mAKAPβ isoform and the remaining predicted human isoforms.

Functional studies

In order to understand the role of mAKAP and the consequence of the identified variant in the HSP pathogenesis, we started by establishing the localization and expression of the protein.

In addition to the experiments with a myc-tagged wild type and mutated construct, we were able to obtain fibroblasts from two patients of this family. For control, we used gender and age-matched wild type fibroblasts obtained from ATCC. We started by checking for potential differences between samples and controls both in protein expression and cellular localization (Figure 2), using a commercial antibody for mAKAP, that regardless of our optimization efforts, presented a band pattern difficult to interpret in western blot. We focused in the bands close to the 250kDa marker, which is the predicted size for the main isoform (mAKAPα in mouse). The band pattern obtained differs among cell types and even among control fibroblasts, which further difficulted the interpretation. By immunocytochemistry, no differences were detected on mAKAP localization between WT and MT isoforms in HEK293T transfected cells using an antibody against the tag (myc). In fibroblasts, we observed different localization according to the permeabilization method used (Figure 2). Using microscopy buffer permeabilization (Harris et al., 2013), both in control and patient fibroblasts, we obtained a similar mAKAP localization to the observed in HEK293T transfected cells using the anti-myc antibody. However, we never observed the perinuclear localization previously reported for mAKAP (Pare et al., 2005).

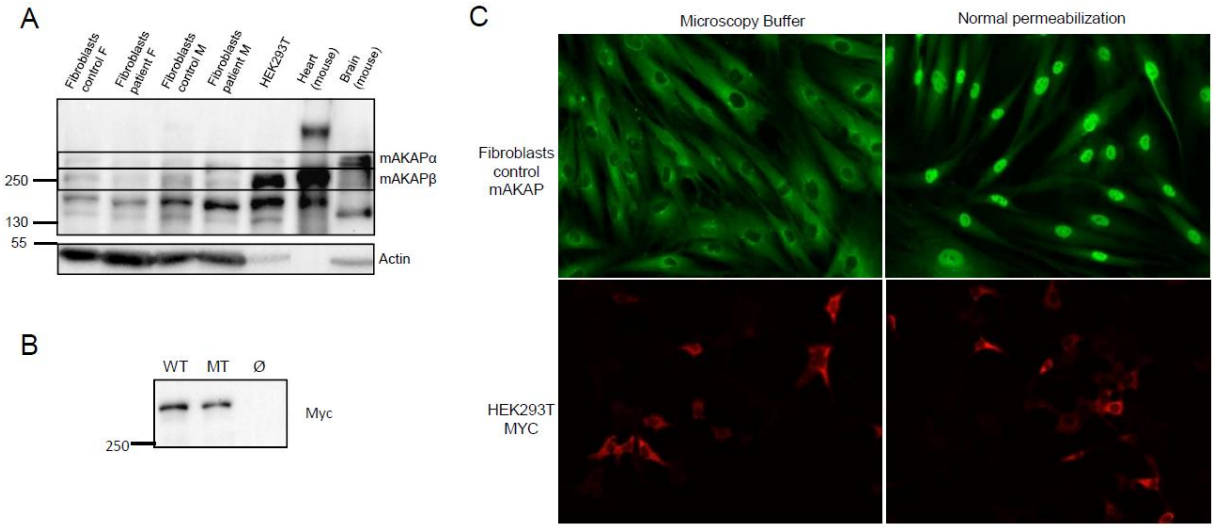


Figure 2. Different mAKAP expression depending on cell type and permeabilization method. (A) Extracts from fibroblasts, HEK293T cells, and extracts from the mouse heart and brain tissues. No major differences were identified between control and patients' fibroblasts, which have a low mAKAP expression. (B) HEK293T cells transfected with wild type and mutant mAKAP plasmid. (C) mAKAP different subcellular localization, in fibroblasts, according to the permeabilization method (microscopy buffer or 0.3% Triton X-100). HEK293T cells transfected with myc-tagged human AKAP6 show expression with a similar localization to the obtained with microscopy buffer permeabilization in fibroblasts.

There are two main mAKAP isoforms in mouse, one corresponding to the full length protein and another corresponding to the translation starting only on exon 4. Taking this into consideration we have looked at the expression, at the RNA level of both isoforms, using 2

sets of primers: one located in exons 2 and 3 and the other pair in exons 5 and 6, which should detect both isoforms. qPCR results showed different expression levels according to the primer pair used when studying the same patient and control fibroblasts: the longest isoform appeared to be more expressed in patients although when looking at both isoforms (primer pair in exons 5 and 6) there was a reduction of expression in patient fibroblasts (Figure 3). Although these results need to be repeated and further explored, they may indicate a potential effect of our variant in the regulation of expression or stability of the mutated mRNA. Also, as there is a much larger number of isoforms predicted for mAKAP in humans than in mouse, an extended study of the human mAKAP isoforms actually being translated in the different tissues is needed for a better knowledge on this gene expression and function. Also, we could observe that the mRNA expression of mAKAP in fibroblasts seem to be much weaker than in HEK293T and SH-SY5Y cell lines, which could indicate that these cells may not be the best cell model for the functional characterization of the mAKAP variants identified (Figure 3).

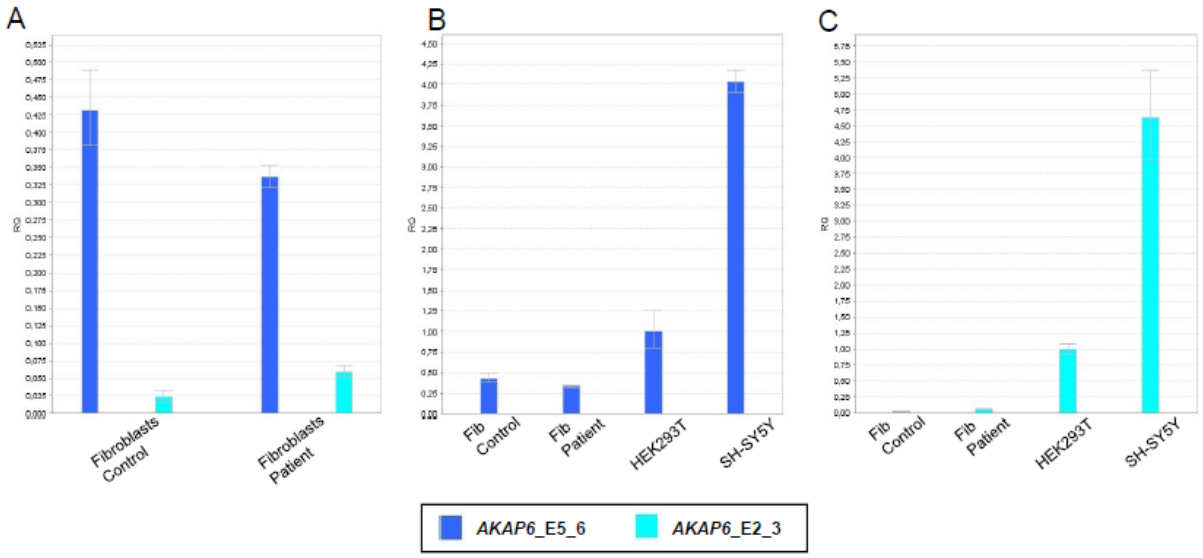


Figure 3. AKAP6 different mRNA expression levels depending on the isoforms and cell types under quantification (n=1). Primer pairs were designed to detect the 2 isoforms described in mice (α and β). AKAP6_E5_6 should inform on the expression of both isoforms and AKAP6_2_3 only on mAKAPα. (A) Different expression according to the primer pair used in control and patient fibroblasts. (B) Different expression of mAKAP in fibroblasts and HEK293T and SH-SY5Y cell lines when using the primer pair on exons 5 and 6. (C) mAKAP expression of different cell lines when using the primer pair localized on exons 2 and 3.

To explore possible differences in protein localization in the different cellular compartments, we performed subcellular protein fractionation (Figure 4), again obtaining different results according to the cell type, namely in fibroblasts and transfected HEK293T cells.

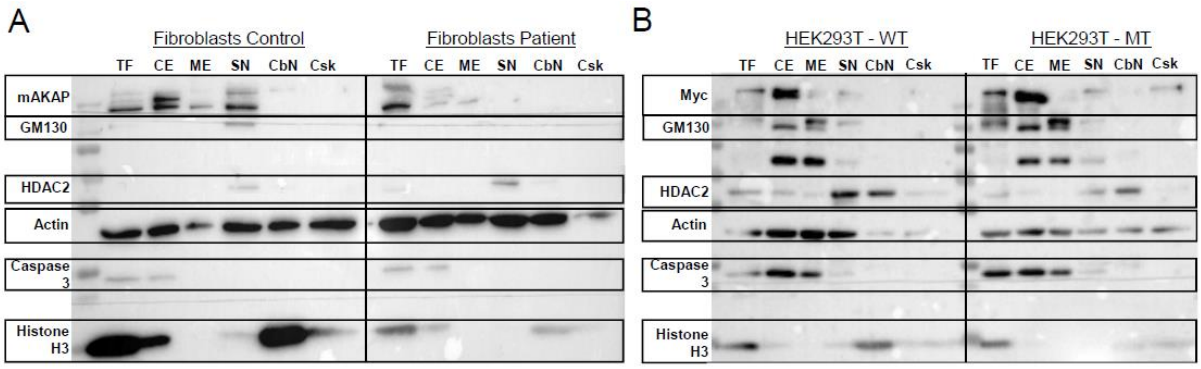


Figure 4. Subcellular distribution of mAKAP in fibroblasts and HEK293T cells (control/patient). TF - Total fraction; CE – Cytoplasmic extract; ME – membrane extract; SN – soluble nuclear extract; CbN – Chromatin-bound nuclear extract; Csk – Cytoskeletal extract.

An extended comparison between patient and control fibroblasts by immunofluorescence labelling using a variety of cellular markers was also performed (Figure 5 and 6) that revealed an interesting difference in the number and size of lysosomes (LAMP1 marker). Patient fibroblasts seemed to have less but larger lysosomes, which appear to be less dispersed in the cell. Minor differences in the mitochondrial network were also observed (TOM20 marker) (Figure 6). To further explore this difference, we studied the two available pairs of patient and control fibroblasts that we differ here by their sex (male and female) and we observed that although this difference is evident in the female pair, when comparing the male pair of control and patient fibroblasts this difference is lost, with the male control fibroblasts having even larger lysosomes than the patient. We also looked at LAMP1 protein expression by western blot and we obtained concurring results (Figure 7), with LAMP1 levels being different among control fibroblasts, resulting in differences according to the pair of control/patient studied.

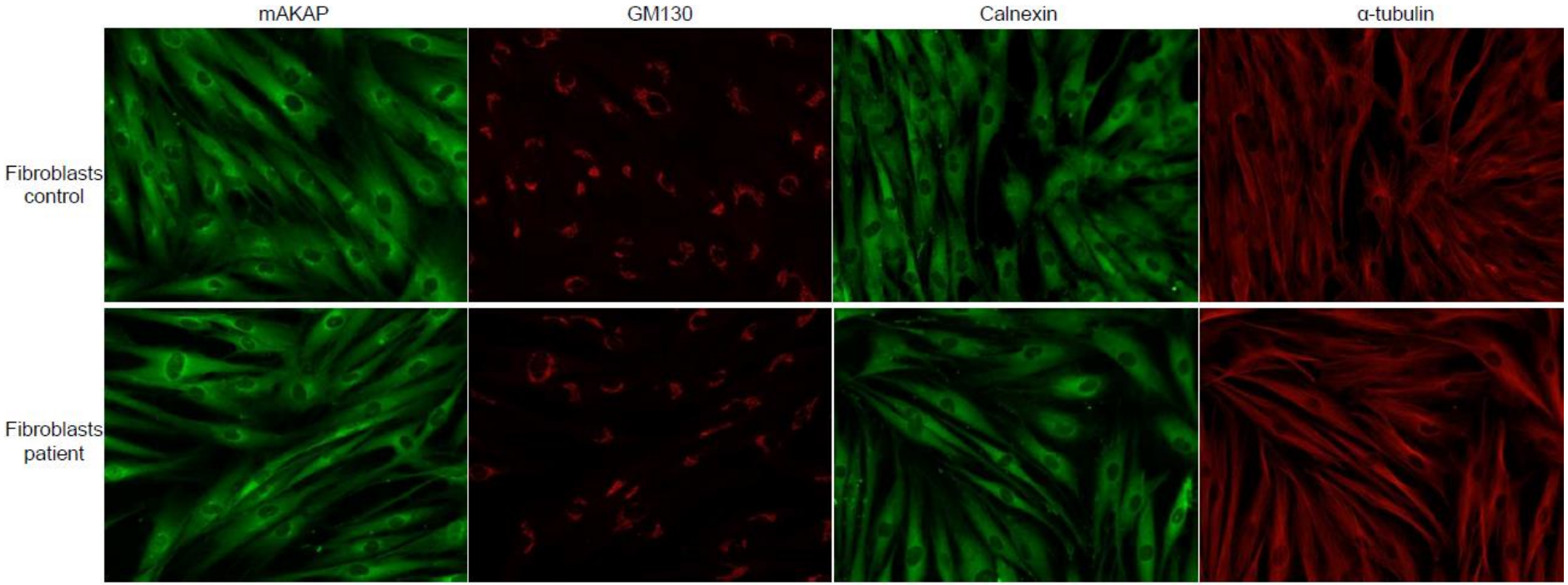


Figure 5. No differences were observed when comparing control and patient fibroblasts for mAKAP and several cellular markers, which included GM130 (Golgi), Calnexin (ER) and α-tubulin (microtubules).

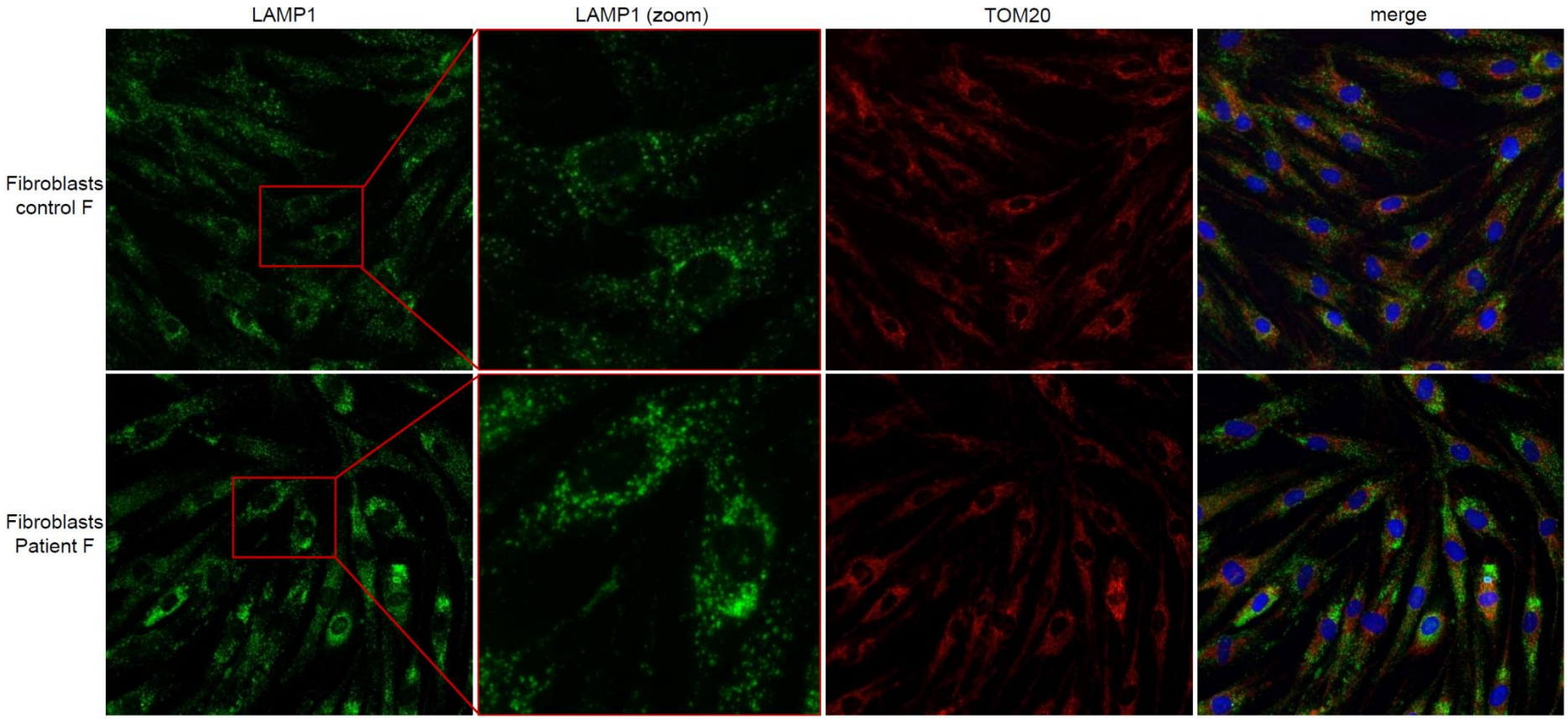


Figure 6. Differences in the number and size of lysosomes (LAMP1) were observed when comparing the pair of female control and patient fibroblasts.

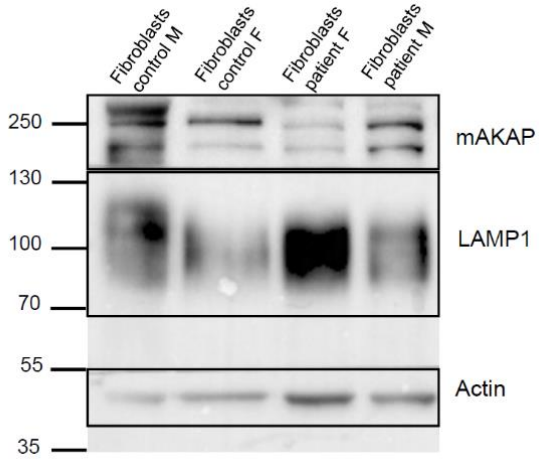


Figure 7. Control fibroblasts seem to differ both in mAKAP and LAMP1 expression, which difficult the comparison between control and patients.

Additionally, as this gene is involved in anchoring the regulatory subunits of protein kinase A (PKA) to the nuclear membrane or sarcoplasmic reticulum, we studied also possible differences in PKA localization in the cell although no differences between control and patient fibroblasts were found (Figure 8).

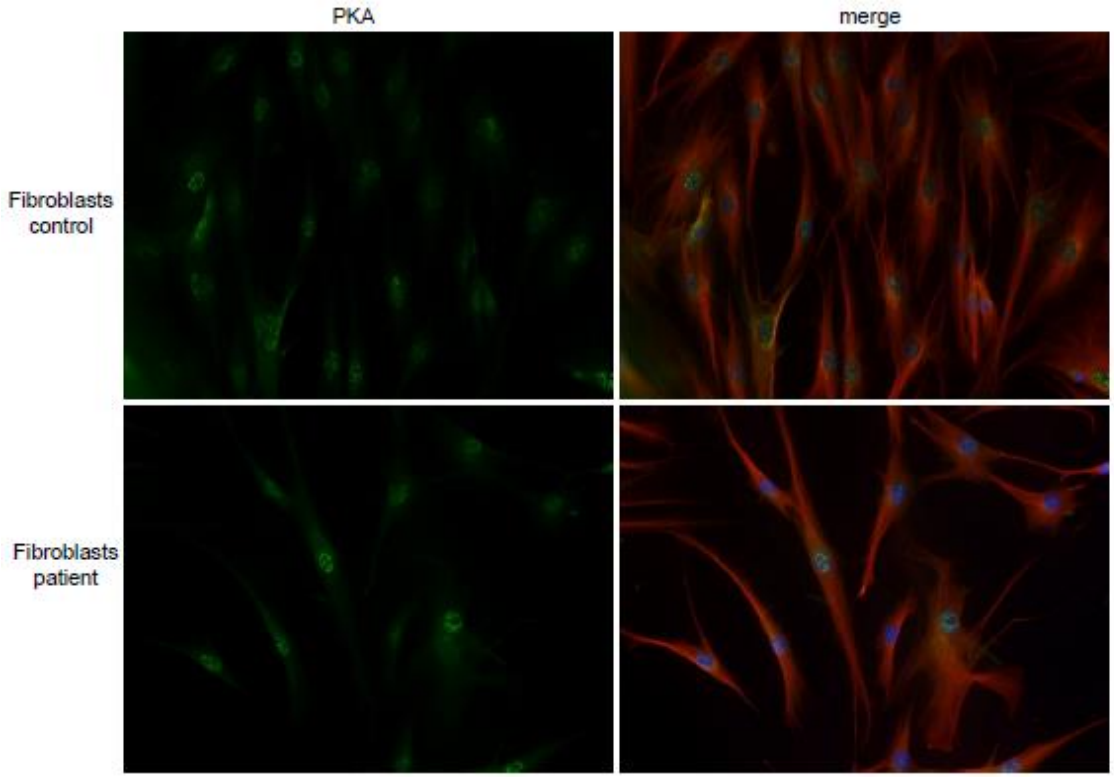


Figure 8. PKA seems to maintain the same localization both in control and patient fibroblasts with a preferential nuclear localization.

mAKAP modeling in zebrafish

To further explore a possible role for mAKAP in HSPs pathogenesis, additional characterization was performed by our collaborators in Paris (J Hazan), using the zebrafish, which proved to be more fruitful. Zebrafish has an AKAP6 orthologue with 2143 amino acids, which has 52% of homology with the human protein. Our collaborators have observed, through in situ hybridization, that AKAP6 is enriched in the zebrafish trunk (probably in the spinal cord and muscles).

To understand the role of this protein in the nervous system of this organism, a morpholino oligonucleotide, directed against the translation initiation site of AKAP6, was injected creating a transient knockdown of the protein. Loss of AKAP6 function led to an abnormal morphology and impairment of locomotion in the larvae. Additionally, a transient disorganization of the muscle fibers and an alteration of the axonal architecture of the primary motoneurons (truncated axons, supernumerary branches and aberrant cell-body outflows) as well as the maintenance of their soma within the spinal cord were also observed. AKAP6 seems to be essential for the axonal migration and growth of primary motor neurons, as well as to intervene in the guidance of the secondary motor neurons. Rescue experiments are now in progress to validate the effect of the variant using injection of WT or MT mRNA.

Discussion

Identification of new disease-causing genes in HSP is very important not only to be able to provide a molecular diagnosis to the families but also for the understanding of the disease pathogenesis. Even with new genes constantly being identified, still many families remain without a molecular diagnosis, probably because there are genes affecting single families. This is the case of the Portuguese cohort which was screened for a large number of HSP causing genes although almost half of the families (49%) remained without molecular diagnosis (Morais et al., 2017).

Exome sequencing led to the discovery of a homozygous variant in *AKAP6* as the most probable cause for this family phenotype. This variant (c.464G>A) is in a conserved position and segregating with the disease, and zebrafish studies confirmed that this gene seems to have an important role in the development of the peripheral nervous system. Rescue experiments in zebrafish using the WT or MT mRNA of AKAP6 are now ongoing that will likely validate the pathogenicity of the variant. Although variants in this gene were never found associated with disease (Passariello et al., 2015) and studies have been focused in the role of mAKAPβ in cardiomyocyte hypertrophy and muscle regeneration in mouse, there is still much to uncover regarding the role of this protein in the human CNS.

In order to understand the role of this protein in humans, we were able to collect fibroblasts from two patients, to explore mAKAP function at endogenous levels, which we considered a better cell model when compared to transfected cell lines where we could only study mAKAP overexpression. However, mAKAP localization in fibroblasts seems to be different from the perinuclear localization described in the literature. We observed that mAKAP localization in the transfected HEK293T cells shows a similar cytoplasmic localization, which was also observed before in COS7 cells with AKAP6 overexpression that could be restored with nesprin-1α cotransfection (Pare et al., 2005). Nesprin-1α is a transmembrane protein that contributes to the targeting of mAKAP to the nuclear envelope of the cardiac myocytes, however this may differ between cell types, as for example in COS-7 cells, where nesprin-1 is not present at considerable levels. Although we consider that these results may be related to the antibody used, we also believe that the results observed may be cell type-specific, as it was suggested by Pare and colleagues that nesprin-1 could be a cell-type selective nuclear envelope receptor for mAKAP (Pare et al., 2005). Also, as AKAP complexes can assemble in a tissue-, compartment- and substrate- specific manner (Dodge-Kafka et al., 2005; Wong and Scott, 2004) and mAKAP seems to be mainly expressed in brain, heart and muscle tissues (Kapiloff et al., 1999), it is possible that different cells may have different mAKAP distribution as well as different isoforms being translated. This could also explain to some degree, the different band patterns obtained in all the protein blots performed, including the ones from the different subcellular extracts obtained from fibroblasts and transfected HEK293T cells. To corroborate this hypothesis, we could observe in the real-time PCR from the control and patient fibroblasts with the primer pairs localized in different exons, that mRNA expression differs among fibroblasts in different directions according to the primer pair used and that AKAP6 mRNA expression is low when compared to HEK293T and SH-SY5Y cell lines.

We hypothesize that, if the human mAKAP isoforms are in fact similar to the mouse ones, our variant may be only affecting mAKAPα as it is present in an exon not coded in mAKAPβ (E3). This could possibly explain the absence of cardiac signs in the phenotype of this family. There are very few studies regarding mAKAPα, but in an interesting study it was found that, mice models lacking mAKAPα in the brain had reduced viability and a smaller body size when compared to WT and heterozygous mice (Michel et al., 2005). Although the mAKAPβ isoform and its role in the heart was the one mainly studied, it seems that mAKAP could be linked to stress responses, as in the case of hypoxia, muscle damage, and retina damage (Kritzer et al., 2014; Lee et al., 2015b; Wang et al., 2015; Wong et al., 2008). In addition to the role of mAKAP in muscle regeneration (Lee et al., 2015b) which could also have some influence in our patients phenotype, it is also described that mAKAPα is required for the survival and axon growth of retinal ganglion cells after injury (Wang et al., 2015). This is in agreement with our collaborator study, in zebrafish, showing that this gene is important for axonal development.

We also considered the observed difference in LAMP1, that revealed the presence of larger and more intense lysosomes in the female patient fibroblasts when compared to the female control, interesting as it was in accordance with what was described for other HSPs like SPG11 and SPG15 (Chang et al., 2014; Renvoisé et al., 2014). However, study of the other control/patient pair was not concordant.

To understand if the disease pathogenesis could be related to disruption of mAKAP binding to PKA leading to an altered localization of PKA in the cell, we checked, in the fibroblasts, a possible alteration in the PKA expression although no alteration was found. As our variant is not present in any of the known locals of interaction already described (Dodge et al., 2001; Passariello et al., 2015) as is the case of the coding polymorphisms studied showing binding alterations (Rababa'H et al., 2013), we believe that the answer probably is in the unique amino-terminal portion of mAKAPα. mAKAP is a major scaffold protein interacting with a wide variety of proteins (Passariello et al., 2015), so it is possible that the unique amino-terminal portion of mAKAPα probably contains enzyme binding sites like PDK1 (Michel et al., 2005) that could allow the assembly of distinct signaling complexes.

Due to the difficulties encountered when characterizing the patients' fibroblasts, since no abnormalities in any subcellular compartment were consistently observed when comparing the 2 affected patients' fibroblasts with controls, we are now focusing in trying to find a better cell model for the AKAP6 study. This may be possible if we could use of the patient fibroblasts to generate induced pluripotent stem cells and differentiate them into neuronal cells.

A more in-depth and vast study of mAKAP and more precisely mAKAPα is needed for a better understanding of this protein role in the human brain and in the HSP pathogenesis.

Table S1. Clinical characteristics of patients from family SR2 with *AKAP6* homozygous variant

Family code	**SR2**		
Individual No. (gender)	II:3 (F)	II:4 (F)	II:5 (M)
Variant	c.464G>A; c.464G>A	c.464G>A; c.464G>A	c.464G>A; c.464G>A
Phenotype	Complex	Complex	Complex
Age at onset in years	6	7	6
Age at examination in years (date)	40 (2006)	35 (2007)	32 (2006)
Disease duration in years	34	28	26
Spasticity LL (0-4)	2	1	2
Spasticity UL (0-4)	0	0	0
Weakness LL (0-5)	4	0	4
Weakness UL (0-5)	4	4	4
Babinski sign	Yes	Yes	Yes
Sphincter disturbances	No	No	No
Dysarthria/Dysphagia	No	No	Yes/No
Cognitive impairment	Yes (mild)	Yes (mild)	Yes (mild)
Other signs	Pes cavus	Pes cavus	Pes cavus and peripheral neuropathy
MRI	Thin corpus callosum, cortical and cerebellar atrophy and pontine dysraphia	NA	Thin corpus callosum, cerebellar and mild cortical atrophy and pontine dysraphia

Abbreviations: M, male; F, female; NA, not available; MRI, magnetic resonance imaging.

For spasticity a scale of 0 to 4 according to the Modified Ashworth Scale was used. Weakness was graded from 0 to 5 using the Medical Research Council Scale for Muscle Strength.

4.2.4. Results from the remaining 9 families studied by whole exome sequencing

Identification of new candidate genes in hereditary spastic paraplegia

Exome analysis of the nine remaining families selected for whole exome sequencing (four families with presumed recessive transmission of the disease and five families with presumed dominant forms) led to the identification of candidate genes in six families with a single candidate gene being found in two of them. Linkage analysis was performed in the five dominant families, in order to reduce the area to focus in, as there was a large number of common variants identified in the exome analysis.

A molecular study of all the potential candidate genes is now necessary to confirm their putative causative role, even more when there are no previous reports that could prove the importance of the proteins encoded by those genes in the human brain. This is also the case and even more difficult, in the families (four in our case) where more than one candidate gene remained after bioinformatics and segregation analysis.

Results

A bioinformatics and segregation analysis of all the filtered variants identified in the exome data from these nine families, allowed the identification of 12 candidate genes (in six families). The candidate genes identified are summarized in table 1.

AR families:

In the four families with presumed recessive transmission of the disease, we found candidate genes in three of them. In family SR13, there was linkage data from a previous screening of 300 autosomal microsatellites spaced 15cM apart (Stevanin et al., 2007b) that identified several *loci* with a maximum LOD score of 1.93 in chromosomes 1, 5, 10, 12, 13 and 18. Therefore, we focused our search in the variants found in these candidate regions. No homozygous variants were found segregating with the disease. As for this family consanguinity was not declared, we focused next on potential compound heterozygous variants, where only one of them was presented in the mother, and we identified only one candidate gene, *SUCO*, with two heterozygous variants segregating perfectly with the disease in this family (Figure 1). This gene is present in the *locus* identified in chromosome 1 and the two variants are rare, with one variant (c. 1811T>C) being a missense variant that causes the alteration of a methionine at position 604 by a threonine (p. Met604Thr) and the other, a small deletion of 4 bases (c. 2110_2111+2del), causing not only a shift of the reading frame but also a predicted loss of the donor splice site.

Table 1. Candidate genes and variants found in six families

Family	Phenotype	Chr	Candidate genes	Accession.Version numbers*	cDNA	Predicted protein effect	State	Nº software **	gnomAD	In favor of a causative role	Against a causative role
A) Families with expected recessive inheritance											
SR13	Complex Early-onset	1	***SUCO***	NM_016227.3	c.1811T>C	p.(Met604Thr)	Het	3/5	-	Gene present in a *locus* identified by linkage analysis, rare variants	
					c.2110_2111+2del	p.(Glu704Valfs*5)	Het	Frameshift, splicing 4/5	1het (0hom)		
SR7	Pure Early-onset	22	*SGSM3*	NM_015705.5	c.1096G>A	p.(Val366Met)	Hom	1/4	30het (0hom)	SGSM family associated with neuronal function	Variant pathogenicity not convincing
SR36	Pure Early-onset	1	***PYCR2***	NM_013328.2	c.431A>G	p.(Glu144Gly)	Hom	2/5	4het (0hom)	Leukodystrophy gene (AR)	Family not large enough for proper segregation analysis
		1	*JMJD4*	NM_023007.2	c.125C>A	p.(Pro42His)	Hom	2/5	16het (0hom)		
B) Families with expected dominant inheritance											
S2	Pure Late-onset	9	*IFT74*	NM_001099222.1	c.326C>A	p.(Thr109Lys)	Het	3/5	38het (0hom)	Candidate gene in ALS-FTD	Segregating compatible only if incomplete penetrance is considered; present on gnomAD
		17	*NPB*	NM_148896.4	c.347C>T	p.(Ser116Phe)	Het	4/4	5het (0hom)		
S18	Pure Late-onset	4	*UCHL1*	NM_004181.4	c.401A>G	p.(Glu134Gly)	Het	3/5	-	SPG79 (AR), PD associated gene	Segregating compatible only if incomplete penetrance is considered or outdated family information
		6	*BMP6*	NM_001718.5	c.211C>G	p.(Leu71Val)	Het	2/4	15het (0hom)		
		6	*HIST1H2AM*	NM_003514.2	c.52C>G	p.(Arg18Gly)	Het	5/5	1het (0hom)		

S57	Pure Early-onset	15	***MAP1A***	NM_002373.5	c.7678C>T	p.(Pro2560Ser)	Het	3/5	-	Present in the candidate list from (Novarino et al., 2014)	Segregating compatible only if incomplete penetrance is considered (a medical reevaluation of the family is necessary)
		2	*ABCA12*	NM_173076.2	c.4740+2T>A	p.?	Het	Splicing 5/5	-		
		15	*SERINC4*	NM_001258031.1	c.854G>A	p.(Arg285His)	Het	2/5	6het (0hom)		

* Exon Number according to Alamut software v2.9.0 that used the NCBI RefSeq numbering scheme (along with simple systematic numbering) on the GRCh37 (hg19) assembly.
** Number of Software predicting a functional effect of all the prediction software for the missense variants; Het, Heterozygous; Hom, Homozygous; Chr, Chromosome; ALS-FTD, Amyotrophic lateral sclerosis – frontotemporal dementia.

In the SR7 family with declared consanguinity, we identified a homozygous variant in *SGSM3* that segregated with the disease in this family (Table 1 and Figure 1). Although this variant pathogenicity is not very convincing as it is only predicted to be pathogenic by Polyphen-2, it is present in an interesting gene as the SGSM gene family has been associated with a neuronal function (Yang et al., 2007). When looking at variants in compound heterozygosity, we found 20 additional candidate genes to be explored. The last family with presumed recessive inheritance where we found candidate variants was family SR36, a smaller family with no declared consanguinity. Interestingly, we found only two homozygous variants segregating with the disease in genes *PYCR2* and *JMJD4*. Both variants are only predicted to be pathogenic by two prediction software and an extended segregation analysis was not possible in this family. However, as *PYCR2* is a gene associated with hypomyelinating leukodystrophy-10 (Nakayama et al., 2015) and the variant that we have identified is located in the pyrroline-5-carboxylate reductase domain, we considered this gene to be the best candidate for this family.

AD families:

The three families with dominant inheritance with potential candidate genes identified (S2, S18 and S57) were analyzed having into consideration the *locus* with maximum LOD score obtained in the linkage analysis (1.2 in families S2 and S18, and 0.9 in family S57). This helped to reduce the number of candidate genes, however, after segregation analysis, in all families, candidate genes could only be causative if we consider incomplete penetrance of the disease. In family S2, both variants from the two candidate genes are present in two family members without disease from the younger generation and both are present in the IV:10 individual. In addition, both were found in several genomes available on gnomAD, although this fact is not enough to discard them as causative if we consider incomplete penetrance and a late-onset disorder. One of these genes, *IFT74,* is also considered a candidate gene for amyotrophic lateral sclerosis – frontotemporal dementia (ALS-FTD) (Momeni et al., 2006). In family S18, three candidate genes were identified that show the same segregation pattern, all of them present on individuals III:7 and III:8 without symptoms of the disease. So, in this case we considered two possibilities: 1) incomplete penetrance or, 2) the two sisters could not present symptoms at the time of the last observation as they were younger than 37 years old and the disease onset seems to be later in the remaining family members. Two of the variants were also found on gnomAD but again as it is a late-onset disorder, they could not be excluded as potentially causative. *UCHL1* was recently associated with HSP (Rydning et al., 2017) being now classified as SPG79, however, it is described as associated with a recessive mode of transmission with two mutated copies being necessary for the disease, which is not the case in our family. Nonetheless, as the number of HSP genes associated with two inheritance

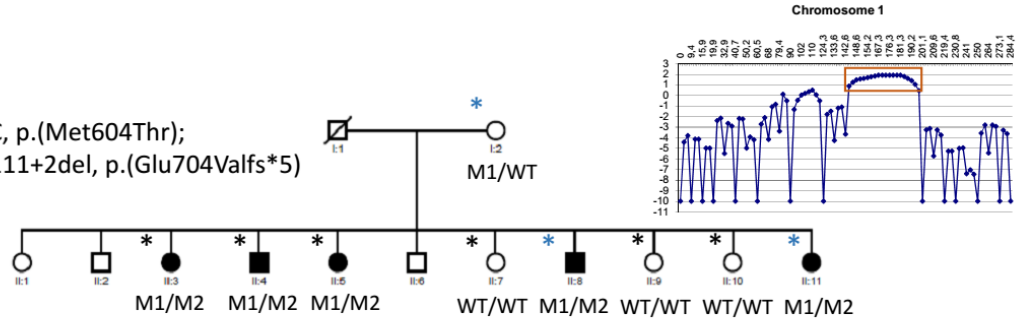
modes such as *KIF1A*, *REEP2, ALDH18A1,* among others is increasing, we should consider and confirm this possibility. In family S57 other three candidate genes were identified, although all of the variants were present in I:2 and III:1 individuals with no disease symptoms. In addition, in this family, five family members showed some pyramidal signs when observed (half-filled squares or circles), and all candidate variants fail to explain the disease symptoms if considered these individuals as affected. A medical reevaluation of the family is necessary to better ascertain the importance of these findings. Interestingly, *MAP1A* was one of the genes identified as a putative candidate in the HSPome generated by Novarino and colleagues that lists 589 candidate genes identified through network propagation of the expanded global network using the set of HSP associated and candidate proteins (Novarino et al., 2014).

In the three remaining families studied, one family with presumed recessive and two with presumed dominant inheritance (SR70, S24 and S91), no candidate genes remained after segregation analysis.

SR13

SUCO

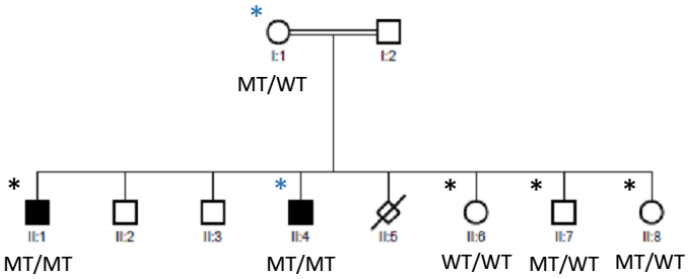
M1: c.1811T>C, p.(Met604Thr);
M2: c.2110_2111+2del, p.(Glu704Valfs*5)



SR7

SGSM3

MT: c.1096G>A, p.(Val366Met)



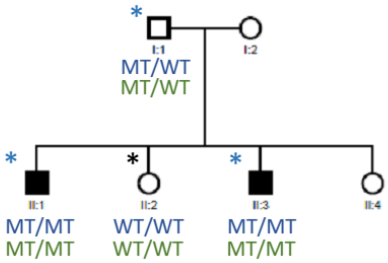
SR36

PYCR2

MT: c.431A>G, p.(Glu144Gly)

JMJD4

MT: c.125C>A, p.(Pro42His)



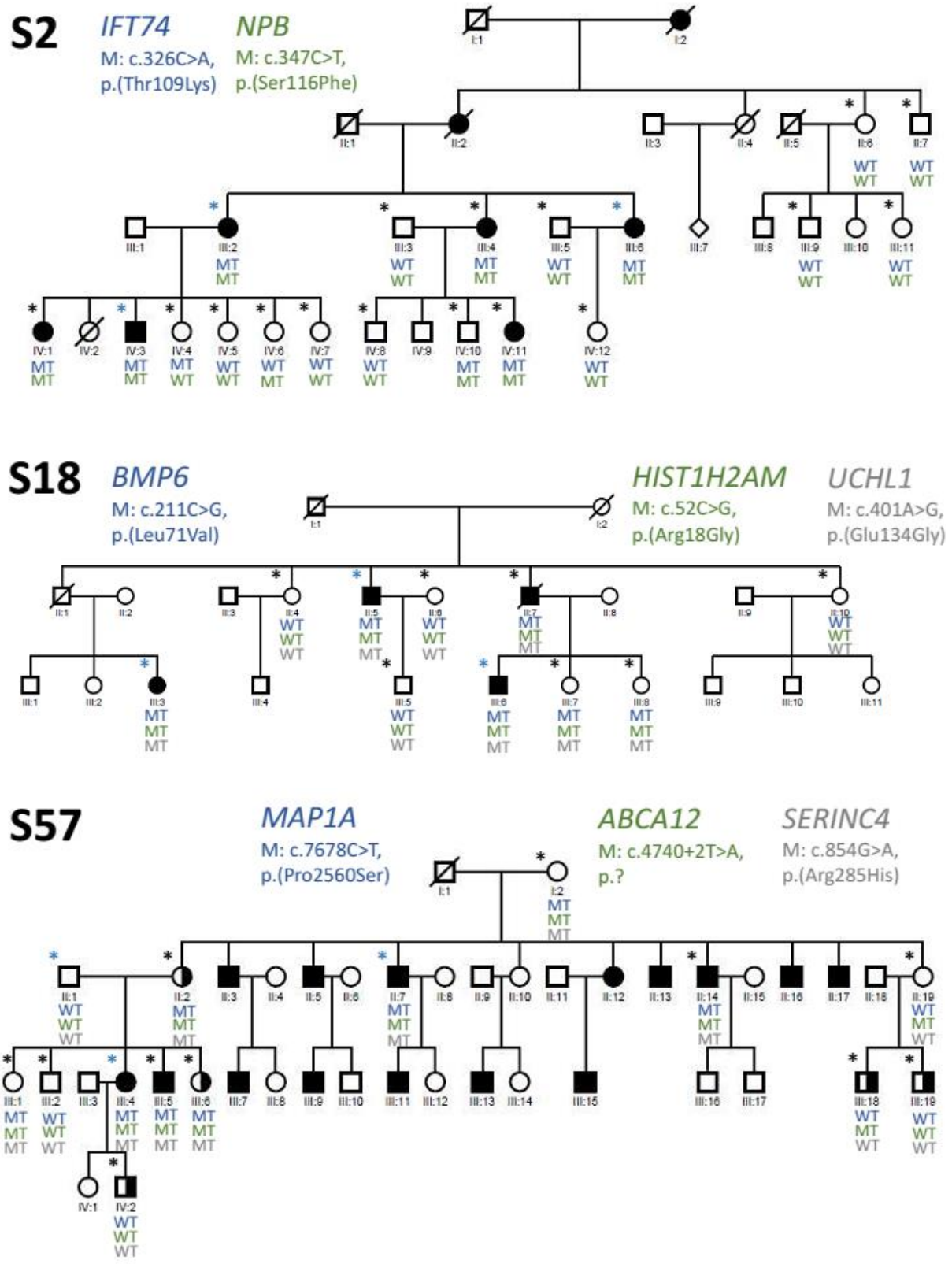


Figure 1. Segregation analysis of the candidate variants identified in the exome analysis of six families. A black circles and squares indicate affected individuals with HSP; half-filled circles and squares on family S57 indicates individuals with minor pyramidal signs. An asterisk in a symbol indicates family members with available DNA and a blue asterisk indicates the individuals selected for exome study. In family SR13 is also shown the locus identified by linkage analysis on chromosome 1 where the *SUCO* gene is present.

Discussion

The identification of candidate genes for HSP in 66,7% (6/9) of the families studied or in 75% (9/12), if we consider also the families with variants in the *AIMP1*, *AKAP6* and *ALDH18A1* genes previously described, is very high and proves that exome sequencing, especially when combined with linkage analysis, is a good strategy for the identification of new HSP disease-causing genes.

In three of the families, no candidate genes remained after segregation analysis, which could be due to several possibilities, the variant could be: 1) present in a low coverage region, 2) a large rearrangement, as the coverage data was not yet analyzed for the exome data or 3) present in an intronic region. Also, as some of these are large families we should also have to consider the possibility of different branches of the family having different disease-causing genes, as discovered before in other HSP large families. It is also very important the revision of the familial data by clinicians because minimal errors in the affected status could lead to an incorrect conclusion on variants segregation.

In the new candidate genes identified in these nine families we found another example of a case that could be related with the clinical overlap between some diseases, with genes being associated with more than one disorder. This is recurrently the case between hereditary spastic paraplegia and leukodystrophy associated genes, where we add *AIMP1* to this list and now also potentially with *PYCR2* variants. We may also have found a new inheritance mode considering the heterozygous variant identified in *UCHL1*, which is not uncommon, as shown for *ALDH18A1, KIF1A and BICD2* genes, among others. These previous associations with HSP or other neurologic disorders help to corroborate the relevance of the finding. However, in the cases where additional candidate genes were found as in family SR36, showing a large difference between the reported and the now observed phenotype (ex. absence of microcephaly), a functional study of the variants identified is needed.

For the remaining candidate genes identified, there are no previous reports that point to an association between the gene and any differential diagnosis of HSP or, showing the importance of the proteins in question in the human brain, potentially explaining the likely pathological mechanisms responsible for the disease, increasing the need to establish the causative role of these variants and genes with detailed functional studies.

The SR13 candidate gene, *SUCO* (SUN Domain Containing Ossification Factor), was first known as *C1orf9* (Røsok et al., 2000) and later as *OPT* (osteopotentia) due to its known role in osteoblastic bone formation, with mouse models lacking *SUCO* developing skeletal defects and potentially neurological defects (Sohaskey et al., 2010). Recently, a new study associated heterozygous *SUCO* variants with mesial temporal lobe epilepsy, a neurological disorder (Sha et al., 2015). In this study it is show that knockdown of SUCO in primary neurons leads to a

striking reduced dendritic length which suggests an involvement of this gene in neuronal development (Sha et al., 2015). Additionally, the subcellular localization of the SUCO protein in the rough endoplasmic reticulum, an important organelle involved in protein synthesis and modification and already related to the pathophysiological mechanism of HSP, could indicate a potential involvement of this protein in the basis of the spastic paraplegia disorders. In family SR13 we have a different phenotype with a complex early-onset spastic paraplegia with no declared epilepsy episodes but with cognitive impairment and kyphoscoliosis.

SGSM3 (Small G Protein Signaling Modulator 3) was the only candidate gene after segregation analysis in family SR7. This gene was formerly known as MAP (Merlin-Associated Protein) due to its association to the Merlin protein encoded by the *NF2* gene, responsible for neurofibromatosis type 2 (Lee et al., 2004). *SGSM3* is ubiquitously expressed and belongs to a protein family (together with *SGSM1* and *SGSM2*) that appears to be associated with the modulation of the small G protein-mediated signal transduction pathway. These proteins are suggested modulators of the small G protein RAP and RAB, which are known to mediate neuronal signal transduction and vesicular transportation pathways (Yang et al., 2007). As these proteins have mouse orthologs expressed in neurons of the central nervous system, a neuronal role was hypothesized for these proteins, which is most evident in SGSM1 that is mainly expressed in brain (Yang et al., 2007). Recently, it was reported that this protein, although in the heart, interacts with connexin 43, being important for its lysosomal degradation (Lee et al., 2017a). This is even more relevant if we consider that connexin 43 is encoded by the *GJA1* (gap junction protein α-1) gene, where alterations were identified in a family with progressive spastic paraplegia (Furuta et al., 2012).

In family SR36, *PYCR2* (Pyrroline-5-Carboxylate Reductase 2) seems to be the best candidate as it was already associated with a hypomyelinating leukodystrophy (Nakayama et al., 2015), which is the case of several HSP associated genes. This gene encodes a mitochondrial protein that catalyzes the last step in proline biosynthesis and seem to be involved in the protection of cells from oxidative stress (Kuo et al., 2016). Although the variant is present in the pyrroline-5-carboxylate reductase domain, the phenotype observed is different from the already described for variants in this gene. This family has a pure spastic paraplegia and no data suggesting microcephaly, present in the families described with *PYCR2* variants, however that could be related to the nature of the variants. In the reported families and in family SR36, the early age at onset is consistent. Regarding JMJD4 protein, there are still few reports that aid in the understanding of its function, being though that it works as an epigenetic regulator, with its knockdown in mouse having no effect in embryogenesis (Hu and Imbalzano, 2016; Yoo et al., 2016). In this family, a plasma amino acid profile was requested that could potentially help to confirm the molecular diagnosis if low proline levels are detected.

The candidate genes in family S2 are *IFT74* and *NPB*. The IFT74 (Intraflagellar Transport 74) protein together with IFT81 form a tubulin-binding module necessary for the tubulin transport within cilia (Bhogaraju et al., 2013). Variants in this gene were associated with Bardet-Biedl syndrome type 20, a ciliopathy responsible for, in the identified patient, obesity, polydactyly, intellectual disability and hypogonadism (Lindstrand et al., 2016). Additionally, it was also consider as a candidate gene for ALS-FTD with one family identified that presented motor weakness in one of the two patients in addition to the former symptoms (Momeni et al., 2006). As the Portuguese family S2 presents a pure late-onset HSP, the phenotype is not in accordance with the ones described in the literature. The other candidate gene, *NPB* (neuropeptide B) is a neurotransmitter expressed mostly in the *substantia nigra* that together with neuropeptide W and the receptors GPR7 and GPR8, seems to be involved in the regulation of feeding behavior and in the manipulation of pain pathways (Singh and Davenport, 2006). *NPB* variants were never associated with any human disease.

In family S18, *UCHL1* (SPG79) was the stronger candidate gene due to its previous association with HSP (Rydning et al., 2017). UCHL1 (Ubiquitin C-Terminal Hydrolase L1) encodes a hydrolase involved in the ubiquitin-proteasome pathway, which is highly expressed in neurons. The two reports linking *UCHL1* variants to HSP and progressive neurodegeneration however, associate this gene with an early-onset of the disease resulting from a recessive inheritance and loss of protein function (Bilguvar et al., 2013; Rydning et al., 2017). Also, both families present a complex phenotype with some visual loss and other complicating features (different in both families). This is not in accordance with the pure phenotype observed in family S18, which also presented a late age-at-onset. The second candidate gene, *BMP6* (bone morphogenetic protein 6), encodes a ligand of the large transforming growth factor-beta (TGF-β) super family. This protein seems to play a role in the induction of mesenchymal cell differentiation into osteoblasts (Paralkar et al., 2010; Rahman et al., 2015). Although *BMP6* variants were still not associated with any neurodegenerative disorder, the BMP signaling pathway is necessary for a good communication between motor neurons and skeletal muscles and was already related to several neurodegenerative diseases, including hereditary spastic paraplegia (Bayat et al., 2011; Osses and Henríquez, 2014; Tsang et al., 2009). *HIST1H2AM* (Histone Cluster 1 H2A Family Member M) encodes a histone, which is responsible for the nucleosome structure, important for the compaction of chromatin, and there are no disorders associated with variants in this gene.

Lastly, in the family S57 we consider *MAP1A* (Microtubule Associated Protein 1A) to be the most convincing candidate. This gene encodes a protein thought to be involved in microtubule assembly that is highly expressed in the brain. The MAP1A expression pattern seems to indicate an involvement in the proliferation and differentiation of neuroblasts and in the maturation of the long spinal fiber systems (Fukuyama and Rapoport, 1995; Oudega et al.,

1998). Additionally, it was one of the candidate genes identified in the HSPome generated by Novarino and colleagues (Novarino et al., 2014). *ABCA12* (ATP-Binding Cassette Subfamily A Member 12) encodes a protein that belongs to the subfamily of ABC (ATP-Binding Cassette) transporters expressed in several tissues including fetal brain and skin (Annilo et al., 2002). Variants in this gene were associated, in several families, with autosomal recessive congenital ichthyosis, a rare skin disorder (Akiyama, 2010). *SERINC4* (Serine Incorporator 4) encodes a protein that belongs to a unique protein family that seems to be responsible for the incorporation of serine into membranes and facilitates the synthesis of phosphatidylserine and sphingolipids (Inuzuka et al., 2005), however this is one of the proteins where there is almost no studies so far.

There is still much to uncover regarding these proteins and variants that only a thorough functional study could help to decipher. These studies are very time-consuming and even more complicated in the families (four in our case) where more than one candidate gene was identified. In order to overcome this, the three families with dominant inheritance must be reevaluated, which could help to understand some of the segregation incompatibilities, and the further extension of these families can help to exclude some of the candidate genes. This would also be very important in family SR36, a small family with one family member still not tested. Finding a second family mutated will also be a very important argument that is currently being explored through GeneMatcher exchanges (Sobreira et al., 2015) and through local analyses of additional families (chapter 4.3). We should not exclude however, that some of these candidate genes might not explain the phenotype and that in some of these families, like in the families where no candidate genes were identified by this approach, the disease cause may be located elsewhere. Our results also show, that most HSP genes are rare affecting only a few families, which indicates that a large number of genes is yet to be identified in this group of disorders.

4.3 Search for additional families with causative variants in the new candidate genes

In order to study some of the new candidate genes identified in the laboratory, a multiplex assay for next generation sequencing was designed that covered 19 genes. It included twelve candidate genes (one of which being *AKAP6),* and seven recently published HSP or HSP-related genes.

The families selected for this study were the families where no molecular diagnosis was found by NGS using the large 70 genes panel and, that were not evaluated by exome sequencing. So, for this mini panel were selected 67 Portuguese HSP families and 125 families from the SPATAX cohort. Some of this work was included in the paper from Marie Coutelier (ALDH18A1 results – presented in 4.2.1). The remaining results from this mini panel are presented in 4.3.1. Additionally, these 67 Portuguese families were also screened, by Sanger sequencing, for all the exonic regions of *AIMP1* with no variants identified and these results were included in the paper submitted (*AIMP1* results – presented in 4.2.2).

4.3.1 Mini panel results

Mini panel sequencing of 19 genes reveals the low frequency of some newly associated hereditary spastic paraplegias genes in the Portuguese population

A multiplex assay for next generation sequencing was designed to study some candidate genes identified in the laboratory and some new HSP genes. In this mini panel 19 genes were included, which encompassed twelve candidate genes (including *AKAP6*) and seven recently published HSP or HSP-associated genes still not included in the HSP panel: *ALDH18A1*, *CAPN1*, *DNM2*, *FARS2*, *GCH1*, *KLC4* and *TUBB4A*.

The families selected for this study were HSP families with no molecular diagnosis found after being studied for a large panel of 70 HSP-associated genes, which included 67 Portuguese HSP families and 125 families from the SPATAX cohort.

The analysis of the Portuguese families studied allowed us to provide a molecular diagnosis to two additional families (2/67) with hereditary spastic paraplegia and identify variants of unknown significance (VUS) in seven families when analyzing the seven new HSP-associated genes and *AKAP6*. Additionally, some VUS were identified in the remaining candidate genes studied but its pathogenic role could not be established or excluded at this point. Unfortunately, no additional families with homozygous or compound heterozygous variants in *AKAP6* were found in the 192 families studied.

Results

Massive parallel sequencing and bioinformatics analysis of nineteen genes associated with HSP (candidate or recently published genes) in 67 index cases of Portuguese families led to the identification of two likely disease-causing variants that could explain the phenotype in 2,98% (2/67) of the families with HSP.

The mean coverage in these 67 index cases was 245 with only 87% of the regions with a coverage ≥ 30x. This low number was due to some amplicons that were not successfully amplified and were then not covered. This also resulted in a large deviation, with the coverage ranging from 0 to more than 2000 reads in some regions. Due to the large number of samples and amplicons, a screening, by Sanger sequencing, of the low coverage regions was not done. After filtering, all the variants considered relevant (rare, predicted as potentially pathogenic and expected to have an impact given the suspected mode of inheritance) were confirmed by Sanger sequencing and segregation analysis was performed in the available family members. An in-house index of coverage was calculated for each exon and for each patient to search for

genomic rearrangements but no large rearrangements that could explain the phenotype were identified.

Novel HSP genes:

In the eight genes selected for analysis, which included *AKAP6* and the seven genes recently reported (*ALDH18A1*, *CAPN1*, *DNM2*, *FARS2*, *GCH1*, *KLC4* and *TUBB4A*), we filtered all the variants with good quality, predicted to cause a change at the protein level (nonsense, amino acid change or a splicing effect), with minor allele frequency inferior to 1% and predicted to be pathogenic by at least one prediction software (SpliceSiteFinder, MaxEntScan, NNSPLICE, GeneSplicer, Human Splicing finder, Polyphen-2, SIFT, MutationTaster, Align GVGD and UMD-Predictor). This allowed the identification of nine variants that could potentially be disease-causing, of which two are likely responsible for the phenotype in two families. Of these two variants, we considered one as disease-causing, as it was a homozygous nonsense variant causing an early stop codon in the *CAPN1* gene in family SR37 (Table 1, Figure 1), which was present in the peptidase C2, calpain, catalytic domain of the protein and segregated with the disease in this family. The second variant identified in *TUBB4A* (family S44) (Table 1, Figure 1), was consider as a likely disease-causing variant, because although segregation analysis could not be performed (due to the absence of additional family members to study), it was a missense variant present in a conserved region, in the tubulin domain of the protein that was predicted to be pathogenic by all the prediction software used and in a gene where a dominant mode of inheritance is expected. Additionally, seven VUS were identified (Table 2), five were heterozygous variants present in genes with presumed recessive mode of inheritance and two were variants present in genes where a dominant mode of inheritance was already described but that its pathogenicity was not very convincing due to discordant predictions on pathogenicity by software used or due to its presence in the public databases, which causative role could only be ascertained with functional studies.

Candidate genes:

In the remaining 11 candidate genes studied, 20 VUS were found in 6 of the genes but no inference of their pathogenicity could be drawn now due to the lack of genetic or biological data.

So, with this panel we were able to provide a molecular diagnosis to two additional families being the yield of diagnosis of 4,3% (1/23) in the families with presumed recessive inheritance, 3% (1/33) in the dominant forms and of 0% (0/11) in the isolated cases.

Table 1. Disease-causing and likely disease-causing variants found in two HSP families

Family code	Phenotype	Inheritance mode [a]	Mutated gene	Accession .Version Numbers	Exon [b]	cDNA	Predicted consequence	State	Nº software [c]	Segregation analysis	Reference
(a) Disease-causing variants											
SR37	Complex Late-onset	Recessive	*CAPN1*	NM_001198868.1	E11	c.1176G>A	p.(Trp392*)	Hom	-	Yes	This study
(b) Likely disease-causing variants											
S44	Pure Early-onset	Dominant	*TUBB4A*	NM_006087.3	E4	c.557C>T	p.(Thr186Met)	Het	5/5	Not possible	This study

[a] The inheritance mode is the one assumed by pedigree analysis.

[b] Exon Number according to Alamut software v2.9.0 that used the NCBI RefSeq numbering scheme (along with simple systematic numbering) on the GRCh37 (hg19) assembly.

[c] Number of Software predicting a functional effect of all the prediction software for the missense variants; Het, Heterozygous; Hom, Homozygous; NBIA, neurodegeneration with brain accumulation.

Table 2. Variants of unknown significance (VUS) of the seven published HSP genes and *AKAP6* identified in seven families

Family code	Chr	Genomic coordinate (GRCh37)	Gene	Accession. Version Numbers	Exon *	Coding change	Predicted change	State	MAF EA	Nº software **	Reason for exclusion
S8	14	33014499	*AKAP6*	NM_004274.4	E4	c.640A>C	p.(Asn214His)	Het	-	4/5	Heterozygous variant of a recessive gene
CI1	14	33293213	*AKAP6*	NM_004274.4	E13	c.6194A>T	p.(Asp2065Val)	Het	-	5/5	Heterozygous variant of a recessive gene
S82	10	97393363	*ALDH18A1*	NM_002860.3	E6	c.602G>A	p.(Arg201Gln)	Het	-	2/5	Inconclusive variant
SR52	11	64953668-64953669	*CAPN1*	NM_001198868.1	E6	c.618-619del	p.(Gly208Glnfs*7)	Het	0.03	-	Heterozygous variant of a recessive gene
S20	11	64953448	*CAPN1*	NM_001198868.1	E5	c.517G>A	p.(Gly173Arg)	Het	0.06	5/5	Heterozygous variant of a recessive gene
SR14	19	10939832	*DNM2*	NM_001005360.2	E19	c.2179C>T	p.(His727Tyr)	Het	0.02	3/5	Dominant with MAF
SR76	6	5369353	*FARS2*	NM_001318872.1	E2	c.550G>A	p.(Asp184Asn)	Het	-	3/4	Heterozygous variant of a recessive gene

* Exon Number according to Alamut software v2.9.0 that used the NCBI RefSeq numbering scheme (along with simple systematic numbering) on the GRCh37 (hg19) assembly.

** Number of Software predicting a functional effect of all the prediction software for the missense and possible splice-causing variants; Chr, Chromosome; MAF EA, Minor Allele Frequency in the European American population from Exome Variant Server (http://evs.gs.washington.edu/EVS/); Het, Heterozygous; Hom, Homozygous

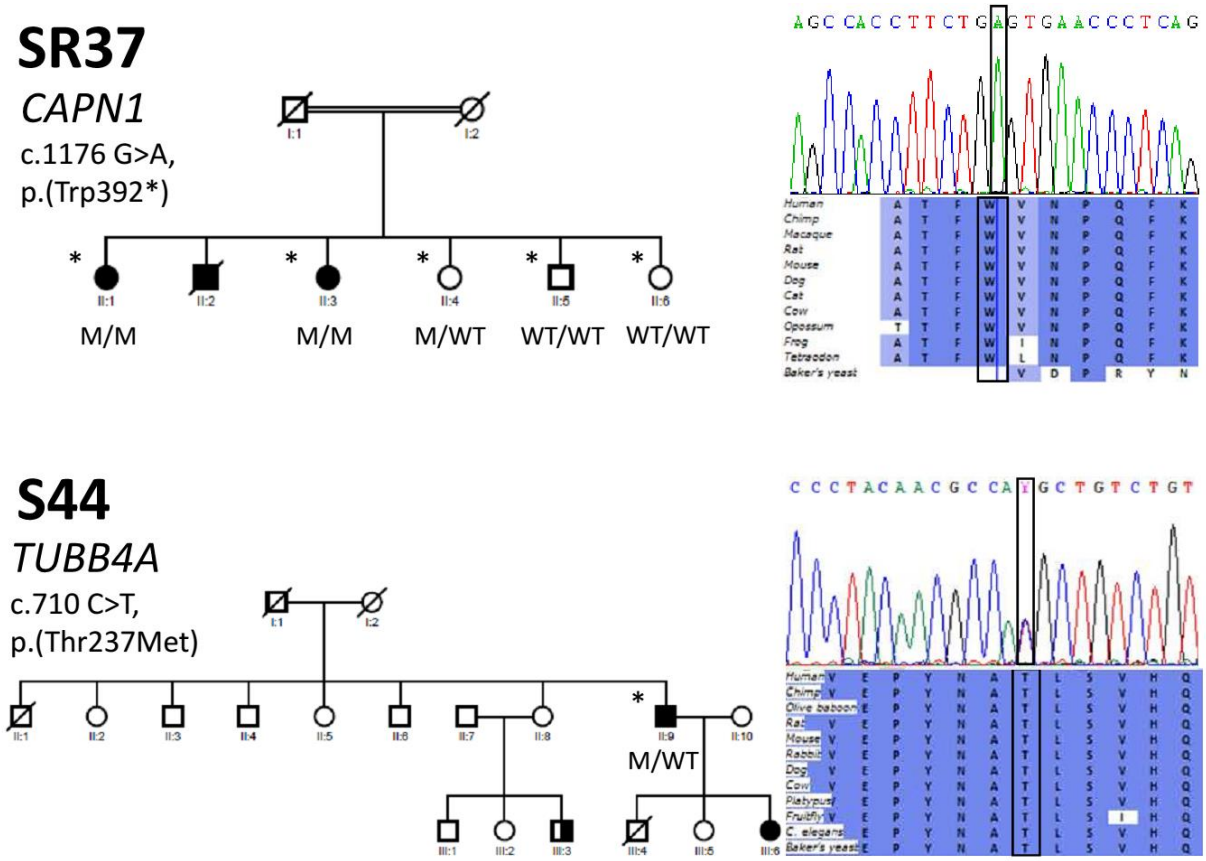


Figure 1. Pedigree with segregation analysis, electropherogram and conservation of the disease-causing and likely disease-causing variants identified in two HSP families. A black circles and squares indicate affected individuals with HSP; an asterisk in a symbol indicates family members with available DNA. Electropherogram and conservation of the region, with the position or amino acid altered marked by a box.

Discussion

The identification of the genetic cause of HSP in only 2,98% (2/67) of the families is not surprising as we are only searching a small pool of genes in families already studied for a large panel of 70 HSP associated genes (Morais et al., 2017). We could not compare this value because there are still no reports on a small panel like this consisting only of newly identified HSP genes. Furthermore, it is in accordance with the findings from the large panel studied (Morais et al., 2017) where it was suggested a low frequency of disease-causing variants in the newly identified HSP causing genes. However, this seem to not be the case for *CAPN1* (SPG76) variants that although only recently associated with HSP (Gan-Or et al., 2016), seems to be less rare with several families already identified that present a phenotype ranging from ataxia to pure or complex spastic paraplegia or a combination of both (Kocoglu et al., 2018; Travaglini et al., 2017; Wang et al., 2016). Clinically, family SR37 presents a complex spastic paraplegia with the presence of cognitive impairment and an age at onset ranging from 28 to 30 years-old, in accordance with the range of phenotypes already described for *CAPN1*

variants, where cognitive impairment was found in patients with a spastic ataxia phenotype (Wang et al., 2016).

CAPN1 (Calpain 1) encodes the large subunit of the ubiquitous enzyme, calpain 1, a calcium-activated cysteine protease widely present in the CNS, where it is critical for the synaptic function (Liu et al., 2008). Knockdown of *CAPN1* in zebrafish models led to a disorganization of the microtubule network in the brain and disruption of the brain development (Gan-Or et al., 2016).

In family S44 we identified an interesting variant in *TUBB4A (*Tubulin Beta 4A Class IVa). This gene was also only recently associated with HSP (Kancheva et al., 2015) but it had already been associated with a large spectrum of phenotypes from autosomal dominant dystonia 4 (DYT4) to hypomyelinating leukodystrophy 6 (Duncan et al., 2017b; Hersheson et al., 2013; Lohmann et al., 2013; Miyatake et al., 2014). Clinically, S44 family showing a pure spastic paraplegia phenotype may expand the phenotype associated with *TUBB4A* variants where only spasticity complicated by other neurologic features was described. However, as no MRI was available from this family, we could not exclude the presence of brain hypomyelination, which was present in all the reported families.

TUBB4A encodes a brain-specific member of the β-tubulin family, highly expressed in the cerebellum and white matter (Hersheson et al., 2013). β-tubulin dimerizes with α-tubulin in a neighboring αβ heterodimer to form tubulin, the main constituent of microtubules, a major cytoskeletal component. Disease-causing variants in other tubulin genes (*TUBA1A*, *TUBB2B* and *TUBB3*) are reported to cause a spectrum of neurologic disorders resulting from abnormalities in axon guidance and maintenance, neural migration and differentiation, indicating the importance of tubulin in the nervous system (Miyatake et al., 2014). A rat model with a *TUBB4A* variant presents progressive loss of myelin in the brain and dorsal column of the spinal cord with an increased number of oligodendrocytes. Additionally, it was also observed an accumulation of microtubules, which involved the entire cell body and distal processes of oligodendrocytes but was not present in axons (Duncan et al., 2017b).

The remaining variants found were all considered VUS, including the variants found in the remaining 11 candidate genes, which need to be functionally studied to conclude on their pathogenicity. Also, as there were some regions not covered, we cannot exclude that some variants located in these regions may explain the phenotype. So, it would be interesting to, in a further study, screen the regions with less than 20X coverage, by Sanger sequencing, mainly in the cases in which only one heterozygous potentially disease-causing variant was identified. A small study like this is interesting as it allows the identification of additional families with disease causing-variants in candidate genes that could be relevant for the HSP pathophysiology. Moreover, it allowed the study of some of the more recently genes associated with HSP, contributing to a better understanding on the frequency of these new genes. A small

panel as the one here described makes sense when we are looking to a cohort like ours where most of the HSP disease causing genes were already studied and excluded. The fact that still 65 of the families remained without a molecular diagnosis, highlight the importance of the search of new disease-causing genes in these group of disorders.

5. Discussion

5.1 Genetic characterization of the Portuguese cohort with HSP

5.1.1 Overview of the molecular strategies and results

In this study, we had access to one of the largest cohorts of families with HSP that historically helped in the description of this group of diseases (Coutinho et al., 1999). Prior to this work, the screening by Sanger sequencing of the most common genetic causes of HSP resulted in a molecular diagnosis in 37% (59/159) of the families. At this stage, *ATL1*, *SPAST* and *REEP1* were screened in families where a dominant inheritance was presumed, and *SPG11*, *ZFYVE26* and *CYP7B1* in the families with a presumed recessive inheritance.
Our molecular study of this large cohort of 159 families (overview in Figure 1) allowed us to achieve a molecular diagnosis in 52,2% (83/159) of the families (Figure 2A).

A stratified approach:
This study focused in the 100 families that remained without molecular diagnosis after these previous studies (Figure 1). In a first step we performed targeted next generation sequencing in 98 of these families applying one of two panels developed by the SPATAX team at the ICM institute (Paris): a smaller one covering 34 genes and a larger panel covering 70 HSP genes. These panels were based on a capture strategy followed by massive parallel sequencing. The study of the two remaining families immediately started by exome sequencing as a previous linkage study had already identified a candidate *locus* with a maximum LOD score of 3.74 in family SR2 (Stevanin et al., 2007b) and of 4.3 in family SR9, falling outside associated HSP genes. The second strategy used was to extend sequencing to the largest families where no disease-causing variants were identified using the NGS panels. Ten additional families were selected and proceed for exome sequencing, one of which (SR13) also had linkage data with several regions identified with a maximum LOD score of 1.93. Exome sequencing resulted in the identification of hundreds of potential disease-causing variants to classify, mainly in families with a presumed dominant inheritance, where we focused on single heterozygous likely disease-causing variants. To reduce the number of variants to focus on, a genome-wide genotyping and linkage analysis was performed in the five families with dominant inheritance of the disease. Finally, all families without molecular diagnosis that were not selected for exome sequencing analysis (with exception of SR26), were studied using a third NGS panel. This smaller panel covered only 19 genes and was designed to study some candidate genes identified in the laboratory, which included *AKAP6*, and some recently published HSP-causing genes. It used an amplicon strategy followed by massive parallel sequencing.
Our first strategy allowed us to provide a molecular diagnosis to 20 HSP families of the 98 studied, but additionally, resulted in the identification of variants of unknown significance (VUS)

in 37 families. This illustrates the difficulties we face in HSP molecular diagnosis and in human genetics in general as some of them may be disease-causing but we lack enough evidence. One of these variants was identified in family SR26 and consisted of a homozygous variant in the *WDR48* gene, initially considered as disease causing, reason why this family was not studied in the third NGS panel. However, as there was not enough evidence supporting the pathogenicity of this variant as it is present in the public databases and is predicted to be pathogenic by only one prediction software, it was reclassified as a variant of unknown significance.

Exome sequencing performed in a total of twelve selected families established the molecular diagnosis in two families up to now but, most importantly, pinpointed candidate genes in 7 additional families, which after functional studies, could be proven as disease-causing and hopefully lead to molecular diagnosis in additional families. Lastly, the study of the third NGS panel that analyzed 67 Portuguese HSP families, allowed the identification of the molecular diagnosis in two families.

<u>An improvement of the diagnosis rate:</u>

Our study of 100 Portuguese families with spastic paraplegia through next generation sequencing established the molecular diagnosis of 24 families (24%), expanded the clinical spectrum of *AIMP1* variants to the HSPs and identified candidate genes in 7 additional families among which are two very strong candidates (*AKAP6* and *SUCO*) supported by the robust genetic and functional evidence. This study, allowed us to increase the number of Portuguese families with a molecular diagnosis from 37 to 52% (83/159), identify potential candidate genes to 7 additional families (4,4%) and detect variants of unknown significance that may potentially explain or help to explain the phenotype in 41 of the 69 families without molecular diagnosis (Figure 2B). Only 17,6% of the families (28/159) remained without any proven or potential genetic cause identified after all strategies applied.

The targeted next generation sequencing performed with the three designed panels, covered in total 74 HSP disease-causing genes if we consider that only 67 of the 70 genes studied with the large panel were HSP associated genes, and, in the small panel covering 19 genes, only 7 were HSP-causing genes already reported. This strategy was a more cost-effective approach to reach the molecular diagnosis in 22 HSP families when compared to exome sequencing, which proved not to be very useful in small families that could hamper the posterior segregation studies. However, exome sequencing is a more in-depth strategy, which allows not only the search for variants in all HSP-disease causing genes but also the identification of potential new genetic causes for the disease. It is also very effective with 75% (9/12) of the families studied with a potential molecular diagnosis.

Figure 1. Schematic illustration of the full cohort of 159 Portuguese families with HSP and overview of the genetic strategies used and mutated genes identified. Our study was focused in the 100 families without diagnosis after screening by Sanger sequencing the most common genetic causes. Genetic strategies are presented in orange boxes, established molecular diagnosis are showed in green and new candidate genes are shown in yellow.

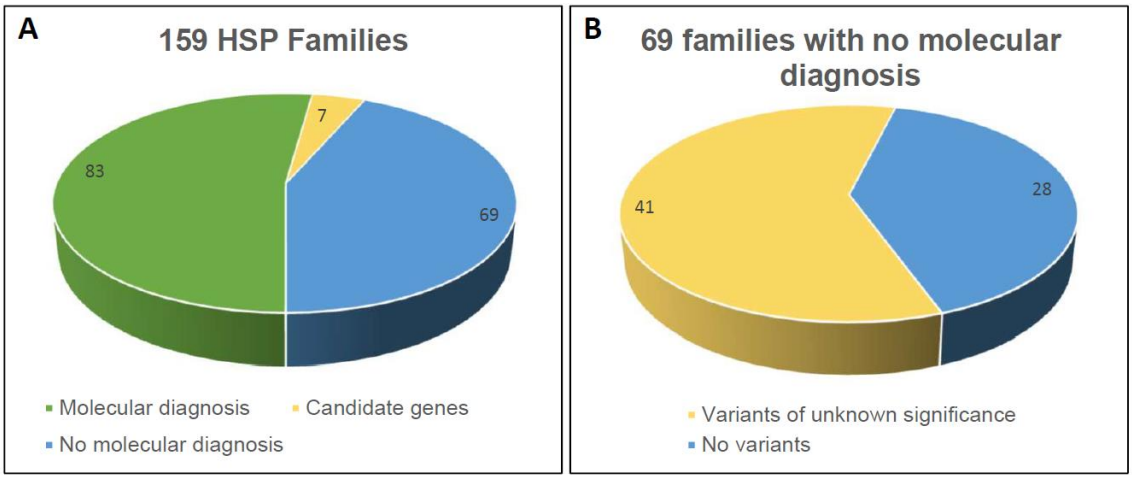


Figure 2. Distribution (by number) of the HSP families according to the presence/absence of molecular findings. A) Distribution of the total cohort of HSP families by identification of molecular diagnosis, candidate genes or no potential diagnosis. B) Distribution of the 69 families without diagnosis according to the identification or not of variants of unknown significance (VUS) that could potentially explain the familial phenotype.

5.1.2 Gene frequencies in the Portuguese cohort

In this study, three multigene panels for HSP were used: a panel covering 34 genes, a larger panel covering 70 genes and lastly, a smaller panel covering 19 genes. The first panel of 34 genes, included the first published and more common causes of HSP and three genes responsible for overlapping phenotypes (*ALS2*, *SACS* and *SETX*). Later, a more extensive panel was created that added 30 HSP-causing genes, plus 3 genes that could mimic the disorder (*FBXO7*, *GJA1* and *SAMHD1*) and three candidate genes (*AP5B1*, *AP5M1* and *AP5S1*). These three candidate genes are part of the AP-5 complex as *AP5Z1* (SPG48) (Hirst et al., 2013b; Slabicki et al., 2010). The third panel was created mainly to screen additional families for disease-causing variants in some candidate genes identified recently in the laboratory, which included *AKAP6,* and was also used to search for variants in 7 recently reported HSP genes including *ALDH18A1*.

The last two multigene panels allowed us to gather data on the frequency of less studied HSP genes, including some of the more recently reported ones, showing that variants in these genes are uncommon. In the first study, only one (or two if we consider also family SR26 with a homozygous variant in *WDR48*) of the disease-causing variants identified, was found in a gene recently associated with HSP and absent in the first panel, a homozygous variant in *C19orf12*. Although 25 of the index cases were tested only for the smaller panel of 34 genes, the additional 36 genes present in the larger panel could only be hypothetically responsible for the disease in ten of these 25 families, as we identified the disease-causing variant in eight of

them and other seven proceed for exome sequencing and no disease-causing variant was found in any of these 36 genes. The third panel applied also evidenced this fact, with only 2 additional diagnoses being achieved in the 67 families studied. This large difference in the variants distribution among HSP-causing genes is quite noticeable when we look at the 83 established molecular diagnosis in the full Portuguese cohort (Figure 3A). We can observe that in the majority of the Portuguese families with HSP, the disease is caused by disease-causing variants in the *SPAST* (SPG4) gene, responsible for a dominant form of the disease, with a frequency of 33,7% (28/83) and in the *SPG11* gene present in 24,1% (20/83) of the families, causing a recessive form of spastic paraplegia. Interestingly, when looking into the frequencies subdivided by mode of inheritance (Figure 3B-D) we can observe that *SPG11* variants are responsible for 44% of the molecular diagnosis identified in the recessive forms and 50% of the isolated cases. In the dominant forms, 63% of the molecular diagnosis were explained by *SPAST* variants, which were also present in 9% of the isolated cases. Variants in *ATL1* (SPG3A), *CYP7B1* (SPG5A), *SPG7*, *KIF5A* (SPG10) and *SPG15* were also relatively common with frequencies ranging from 3,6% to 6% in the full cohort. We were able to identify, using these panels, some variants in these genes commonly responsible for the disease like *SPAST, SPG11* and *SPG15*, due to the fact that there were a few families not yet screened for variants in these genes. In the case of family SR4, which is in fact an isolated case, in whom we identified a *de novo* variant in the *SPAST* gene, previous studies only included *SPG11* and *SPG15* because a recessive mode of inheritance was assumed due to the absence of family history. The remaining genes identified were responsible for the phenotype in unique families of the Portuguese cohort.

This type of distribution, with the majority of diagnosis being of SPG3A, SPG4 and SPG11, and a smaller percentage being assigned to some of the remaining identified genes are also common findings in other countries like Canada and Greece (Chrestian et al., 2017; Lynch et al., 2016) or even if we take into consideration the SPATAX cohort (Tesson et al., 2015). In the large Canadian cohort, parallel sequencing using a panel of 51 genes or exome sequencing, showed that in addition to the presence of some families with variants in genes like *PLP1* (SPG2) and WASCH5 (SPG8), there were also variants identified in genes that could mimic spastic paraplegia, such as *SLC2A1*, responsible for dystonia among other disorders and *VAMP1,* associated with spastic ataxia (Chrestian et al., 2017).

The total yield of molecular diagnosis of 52% was similar to the one obtained when separating the families by inheritance mode (dominant, recessive and isolated cases), which was of: 53% (32/60) in the families with a presumed recessive mode of inheritance, 51% (39/76) in the dominant families and 52% (12/23) of the isolated cases. This high percentage of molecular diagnosis in sporadic cases show the importance of studying these cases, which are frequent in clinical practice.

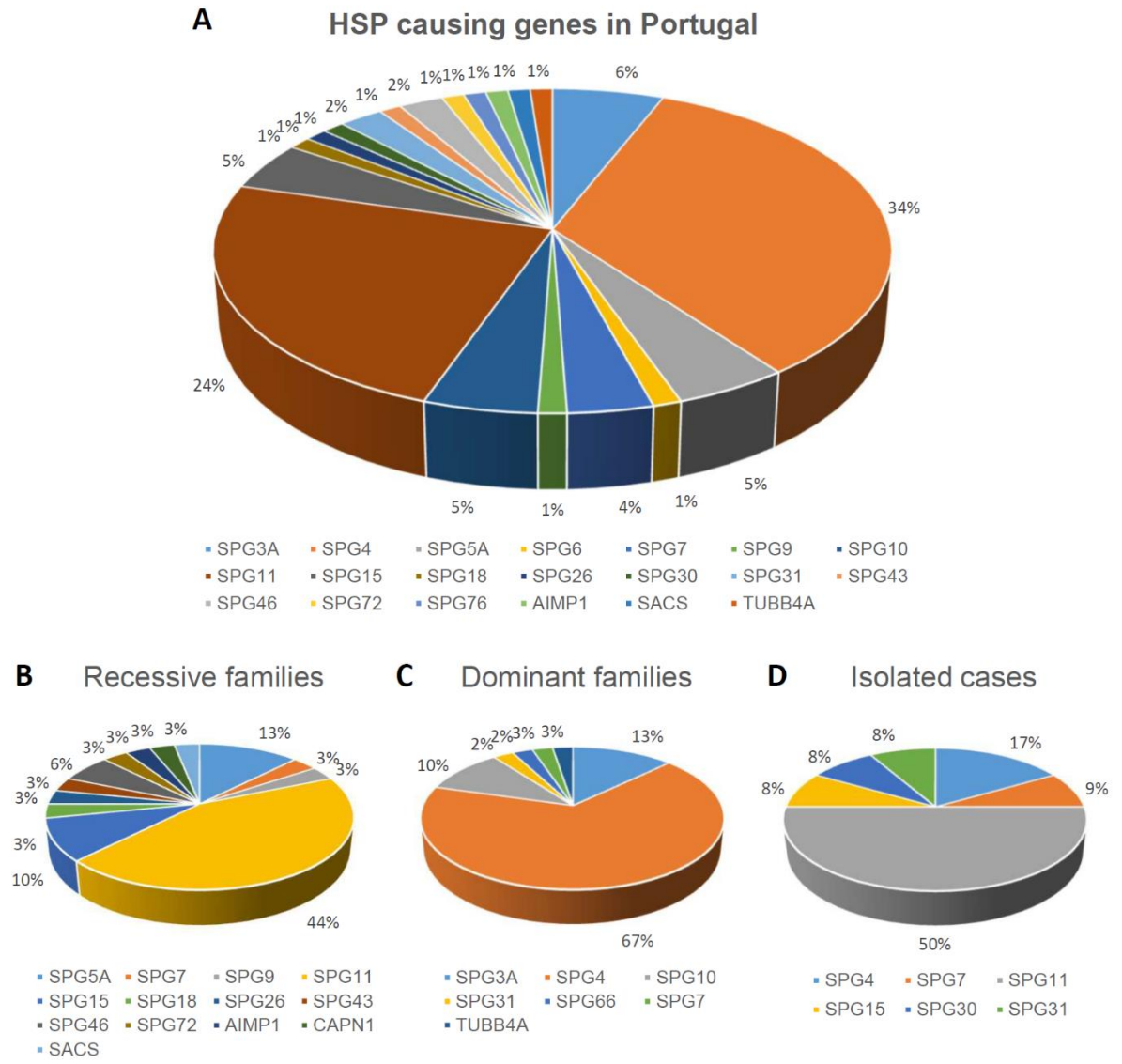


Figure 3. Distribution (by frequency) of the disease-causing variants found in 83 families from the full cohort of Portuguese cases. (A) Full cohort, (B) Recessive families, (C) Dominant families, (D) Isolated cases.

5.1.3 Molecular heterogeneity

This study highlighted the molecular, genetic and clinical heterogeneity between neurodegenerative diseases mostly with the identification of new spastic paraplegia genes like *AIMP1* and *ALDH18A1*, previously linked to other partially overlapping neurodegenerative disorders. In the two initial panels, six genes that could mimic spastic paraplegia were included (*ALS2*, *SACS*, *SAMHD1*, *GJA1*, *FBXO7* and *SETX*) and, although variants were only identified in *SACS*, this only highlights the low frequency of causative variants in these particular genes in the Portuguese population. In fact, genes primarily associated with other neurodegenerative disorders, considered HSP differential diagnosis, are being increasingly found responsible for

purer forms of spastic paraplegia, with their range of expected phenotypes being enlarged. This is the case of *AIMP1, ALDH18A1, CAPN1* and *TUBB4A* genes, also found responsible for the disease in single Portuguese families.

The clinical and molecular bridge between HSP and hypomyelinating leukodystrophies was a common finding in our families not only due to the variants identified in *AIMP1* and *TUBB4A*, but also if we consider the homozygous variant identified in *PYCR2* as the most likely candidate to explain the phenotype present in family SR36. Up to now, at least nine additional genes (*SLC16A2, GJC2, PLP1, FA2H, ABCD1, ADAR, HSPD1, EIF2B5* and *SAMHD1*) are a common link between leukodystrophies and HSP reinforcing this heterogeneity and the possibility of shared disease mechanisms.

The link between spastic paraplegias and ataxias is also evident with genes being, for long, associated with both diseases such as *SPG7* and *PNPLA6*. In our families, in addition to variants in *SPG7* and *SACS*, we also identified variants in *GBA2, CAPN1, ZFYVE26* and *TUBB4A,* traditionally described as responsible for ataxia and spastic paraplegia phenotypes. This association is larger with many more shared genes, phenotypes and consequently pathways between these disorders (Synofzik and Schüle, 2017), which strengthens the causative role of the variant identified in the candidate *UCHL1* as responsible for the phenotype in family S18.

The constant increase in the interconnections between neurodegenerative disorders that share pathways and genes, which in turn show an increasing range of associated phenotypes, is leading to the loss of borders between these disorders. So, nowadays this division makes less sense and we should probably think of them as a complex network of phenotypes, not discarding this possible interconnection particularly when searching for the disease genetic cause.

5.2 Molecular approaches for HSP diagnosis

5.2.1 Unravelling “old” and new spastic paraplegia genes

Establishing a molecular diagnosis in HSP, a group of disorders with more than 70 genes already associated with the disease, has proven to be a challenging task, for which several approaches could be applied.

Ancient diagnosis strategy:

From a long time, most diagnosis were achieved by sequential Sanger sequencing of the most common genetic causes, taking into account the suggested inheritance mode and the familial

phenotype. Considering these, families with an apparent recessive inheritance and with patients presenting a thin corpus callosum should be tested for the presence of variants in *SPG11* and *SPG15;* patients with cerebellar atrophy and/or optic atrophy should be screened for variants in *SPG7;* and when a dominant mode of inheritance was presumed, it was advised to test for variants in *ATL1* (SPG3A) and *SPAST* (SPG4) (Depienne et al., 2007; Kumar et al., 2015; Souza et al., 2016). This strategy was effective, to some extent, but the majority of families remained without a molecular diagnosis.

Current diagnosis strategy:

Nowadays, with the sequencing throughput provided by next-generation sequencing, either with gene panels, whole-exome or whole-genome sequencing, a large number of genes can be sequenced at the same time, allowing for a large number of diagnosis and a better notion of the prevalence of, until now, less studied genes. DNA-microarray–based diagnosis is a strategy that could also be applied (Luo et al., 2013) but as it can only be used to detect common point mutations, is not very useful or popular due to the large variety of variants and genes already identified.

In our study, the panel-based strategy was very useful for achieving additional molecular diagnosis, with the study of 89 genes between candidate and reported HSP genes. However, even considering the identification of several interesting variants considered as VUS, still most of the studied families remained without a definitive answer. Exome sequencing, although more expensive, prove to be more efficient with only 3 of the 10 families studied without any candidate identified after analysis. In these three families however, it is still possible that the molecular cause is a large rearrangement that should be studied by the analysis of the coverage data. In addition, the possibility of digenism should also be looked at, which could lead to the identification of new genetic causes.

Future of diagnosis:

The future of diagnosis will probably be genome analysis, as it allows the identification of not only exonic but also deep intronic variants and allows a better detection of large rearrangements, which are common causes of pathogenicity, accounting for around 10% of the variants present in *SPG4* and *SPG11*. In addition, the analysis of these large data sets should evolve in order to assume that multiple inheritance patterns can be associated to one given gene and that digenism could also be involved. The interpretation of VUS are also a challenge in diagnosis and the development of better bioinformatics prediction tools or of biological assays as performed for *CYP2U1* (Durand et al., 2018) are necessary. Nowadays, is important that a molecular geneticist extend its skills both in bioinformatics and cell biology to better answer the daily challenges.

Evolution of diagnosis strategies, panels versus exome/genome sequencing

The fact that almost half of the families remained without diagnosis after such a large study shows the importance of the identification of new spastic paraplegia genes. This study provided the first evidence for *AIMP1* variants as causative of an HSP phenotype, identified *AKAP6* as the most probable genetic cause of SPG32 (Stevanin et al., 2007b) and identified an additional family with *ALDH18A1* variants. There is also convincing evidence for the pathogenicity of the *SUCO* variants and its association with the HSP phenotype. However, both for *AKAP6*, *SUCO* and the variants identified in the remaining families with exome data, only functional studies can prove this association. There are probably much more genes to be identified and new associations to be made but we now aim to prove the association of these new genes with HSP and the pathogenicity of the variants identified, further confirming the already strong genetic evidences.

The search for new genes will become progressively part of the diagnosis strategy and not only a pure research quest. Part of the novel HSP genes are already known as disease-causing. Indeed, as ours and other groups results indicate, many families have variants in genes already associated with other neurodegenerative disorders, such as *AIMP1* and *ALDH18A1,* as responsible for their clinical phenotype. Thus, maybe the most effective approach for the families that remained without molecular diagnosis should be exome sequencing with a posterior analysis of a large, virtual panel including neurodegenerative disorders associated genes. This allows, in addition to the study of HSP genes, the identification of variants in genes associated with related disorders like neuropathies or leukodystrophies, among others. The identification of variants in genes responsible for these other neurodegenerative disorders in HSP patients could result in a better comprehension of the disease mechanisms and facilitate gene discovery, as proven before (Novarino et al., 2014). Additionally, these virtual panels can be extended over time, as more and more genes are being associated with these disorders and also, this type of analysis reduces the risk of the identification of incidental findings, not related to the phenotype in question that occur when analyzing exome or genome data or even mendeliomes, that target all genes previously linked to single-gene disorders. A similar strategy was used to diagnose a cohort of 319 patients with undiagnosed cerebellar ataxia but the analysis was focused only on ataxia and spastic paraplegias associated genes, with a percentage of diagnosis of 28.5% but where a post hoc re-analysis of data with newly identified genes allowed the identification of causative variants in 4 additional families (Coutelier et al., 2018). Moreover, these studies could also be used, using a different analysis, to search for new hereditary spastic paraplegia genes (Ku et al., 2012). The identification of new spastic paraplegia genes is very important not only to provide

the molecular diagnosis to the patients but also for a better understanding of the disease pathogenesis.

5.2.2 Next generation sequencing challenges

The use of next generation sequencing both through multigene panels and exome sequencing, in addition to the twenty-four molecular diagnosis established, highlighted the difficulty in the interpretation of several potential disease-causing variants. This difficulty, particularly for novel missense variants, is likely to occur due to the fact that these interpretations are usually based on pathogenicity predictions from diverse software that commonly are discordant. These software's use for the prediction, sequence data (usually the protein sequence) aligned with homologous sequences and in the presence of an evolutionarily conserved position result in a high score. These positions are analyzed taking into consideration the amino acids present in the multiple alignment and the severity of the amino acid change, where the physiochemical properties of the amino acids are also considered. Some of these prediction software´s could also use structural information, when available, which could help to better ascertain pathogenicity. Even with the overall improvement of these software's over the years, these are only predictions that use different algorithms and which have an associated false positive and false negative error rate (Dong et al., 2015; Ng and Henikoff, 2006). Another factor that should be checked to ascertain a variant pathogenicity is its presence in the public databases, from dbSNP to more recent databases like the Exome Aggregation Consortium (ExAC) and the Genome Aggregation Database (gnomAD) that aggregate sequencing data from a wide variety of large-scale sequencing projects. This information is very useful and allows the exclusion of common variants. However, when looking at variants with a minor allele frequency below 1%, care should be taken and exclude cannot only rely on the presence in public databases, particularly for variants in gene associated with recessive inheritance disorders. Although uncommon, some variants could have a relatively high frequency (but below 1%) and be disease-causing. This is the case of the *SPG7* p.(Ala510Val) variant that is relatively common, with 794 heterozygous and 2 homozygous identified on gnomAD, which was already confirmed as pathogenic by several authors (Elleuch et al., 2006; Klebe et al., 2012; Sánchez-Ferrero et al., 2013). A frequency this high could however be enough for exclusion when looking at variants in genes associated with dominant inheritance diseases. Finally, we should also check the amino acid conservation of the region and if the altered residue is present in a protein domain, which could be relevant for protein function, thus increasing the likelihood of pathogenicity. This difficulty in variant classification and interpretation is even more complex in WES or WGS studies where the gene effect is unknown. In these cases, not only the variant

potential pathogenicity has to be ascertained but also the gene role and is potential effect in the brain has to be looked at.

In this study, several interesting variants that could potentially be disease-causing were identified that, as their pathogenicity could not be proven at the moment, were classified as variants of unknown significance (VUS) or considered as candidate variants (in candidate genes) when resulting from exome data. Functional characterization of the homozygous variant identified in *KIF5A* will be very important for the determination of a molecular diagnosis in this family. This variant segregates with the disease in a recessive manner, although until now only dominant inheritance was reported associated with this gene. To fully ascertain the pathogenicity of this variant, present in the protein kinesin motor domain and predicted to be pathogenic by the five software used, a more in-depth study of this family and variant segregation is needed in addition to the functional characterization. Although there is no suggestion from the coverage data of any large rearrangement in this gene, this possibility should be excluded, before establishing a potential reduced effect of this variant on the protein, potentially explaining a recessive inheritance mode.

The number of genes for which dominant and recessive transmission of the disease is already reported is increasing, which includes *ATL1,* described for a long time as being only inherited in a dominant manner that has now at least one variant reported as segregating in a recessive mode (Khan et al., 2014). The contrary could also occur as it was demonstrated more recently for *ERLIN2* (Rydning et al., 2018).

In addition to the homozygous variant in *WDR48* previously discussed, a variant in *BSCL2* is also a very good candidate for establishing a molecular diagnosis. This variant is present in a gene where a dominant transmission mode was described, is localized in a conserved region, in the same domain as two reported variants and is predicted to affect function by five software. We would like to better study this family because, although this variant was identified also in the non-affected mother, and in a family where a recessive inheritance mode was presumed due to the presence of consanguinity in the family, we believe it could be disease-causing as incomplete penetrance was already observed for this gene (Musacchio et al., 2017; Windpassinger et al., 2004). Two more families presented a single heterozygous variant (c.895G>A) in this gene, although segregation analysis was hampered due to the lack of additional family members. This variant was however less interesting as it was present in the public databases, with 14 heterozygous identified on gnomAD.

Segregation analysis and functional study of the four variants identified in two different families, SR65 and S86, present in genes known to be members or interactors of the adaptor protein complex 5 (AP-5) should also be interesting in order to ascertain the possibility of digenism. Both families show heterozygous variants in two different genes (*AP5B1* and *AP5Z1* or *AP5Z1* and *SPG11*) known to interact at the protein level (Hirst et al., 2013b). Although *AP5B1* is a

candidate gene with still no reports of disease-causing variants leading to spastic paraplegia, it is a member of the AP-5 complex. This complex is the most recently discovered AP complex and, as the other four AP complexes (AP-1 to 4), is comprised of four subunits forming heterotetramers and is involved in vesicle trafficking. AP complexes have distinct localization and functions, with AP-5 seeming to be involved in the traffic of proteins from endosomes to other membranous compartments. Both AP-4 and AP-5 complexes were associated with hereditary spastic paraplegia but although variants in all the four members of the AP-4 complex were already identified in HSP families, only variants in AP5Z1 of the AP-5 complex were found until now (Hirst et al., 2011, 2013a). This could indicate a potential pathogenic role for variants in the remaining three genes encoding AP-5 subunits and their association with spastic paraplegia, which led us to include them as candidate genes. Anyhow, VUS require investments in bioinformatics and cell biology for better biological evidences of pathogenicity (Durand et al., 2018).

In this study, some of the interesting variants found were excluded due to the incompatible segregation analysis and thus not reported. Although considered as not disease causing in these families, we did not report them as likely benign because we did not perform any functional study to confirm that. In addition, we must consider the possibility that the lack of segregation observed could be due to numerous factors such as inaccuracies in the familal information sometimes due to the presence of slightly different disorders, families where there are different genetic causes in different family members, or even sample swaps. So, it is essential that the family information is correct and as much updated as possible, mainly in late-onset disorders, and the presence of incomplete penetrance should also be considered when analyzing variant segregation.

A large study like this, even with all the molecular approaches performed, still was unable to provide a definitive molecular diagnosis to 47,8% of the families. Although some additional diagnosis could be provided after a molecular study of some of the variants identified, we should be aware that some variants present in the genes already tested could have been lost due to the limitations of the technique. NGS is a great method for the sequencing of a large number of amplicons in a fast and cost-effective way, either using panels, exomes or genomes, however it also has some limitations. The presence of variants could be lost in regions poorly or not covered; large rearrangements, although possible to identify them with coverage analysis, could also be difficult in poorly covered regions; inversions could be almost impossible to detect due to the short size of the amplicons tested; and in the case of exomes and gene panels, we could be losing the variants in regions not targeted like introns or promotors. Some other reasons for losing the genetic cause are mainly in the data analysis an interpretation where for example, the potential for digenism should be considered.

5.3 Unravelling spastic paraplegia mechanisms

5.3.1 Pathophysiological mechanisms involved in HSP and axonal degeneration

Hereditary spastic paraplegias are a group of disorders pathologically characterized by axonal degeneration, maximal in the longest tracts that constitute the longest motor and sensory axons of the central nervous system (CNS). HSP is considered a length-dependent 'dying back' axonopathy as the degeneration seems to start in the distal ends of the axons and progress toward the cell body (Crosby and Proukakis, 2002). As the longest corticospinal axons, the ones most affected by this disease, control the lower motor neurons responsible for innervating lower-limb muscles, it explains the main clinical features of HSP (Blackstone, 2018). On the other hand, as in complex forms, the degeneration of other neurons also occur, suggest that this dying back axonopathy may not reflect the degenerative process in all HSP entities (Denora et al., 2016).

The identification of more than 70 genes associated with HSP revealed several common pathophysiological themes that seem to be at the basis of the disease that included defects in: axonal transport, microtubule dynamics, ER morphogenesis, membrane trafficking and recycling, mitochondrial function, lipid metabolism, myelination, autophagy, DNA repair and nucleotide metabolism. These are all defects in basic cellular functions present in most cells, but that seem to primarily affect neurons due to their remarkable morphology. Neurons are highly polarized cells, that in the case of the long neurons affected by HSP, can have axons reaching up to one meter in length, which makes them sensitive to alterations in many cellular processes (Blackstone et al., 2011; Denton et al., 2016). These pathophysiological themes seem however to be common to most neurodegenerative disorders.

Alterations in axonal transport are a common observation in a variety of neurodegenerative diseases, including motor diseases like HSP, peripheral neuropathies and ALS. This intracellular transport of cargo such as proteins, mRNA, lipids, membrane-bound vesicles and organelles, mostly synthesized in the cell body, is essential for their correct spatiotemporal distribution in the axon and for the maintenance of long-distance communication between cell body and synaptic terminals. This transport is mainly based on microtubules, which polarity directs the movement of the molecular motors along them (De Vos and Hafezparast, 2017). Alterations in membrane trafficking and degradation mechanisms like autophagy are also consistent for most neurodegenerative disorders (Wang et al., 2013). The proper functioning of membrane trafficking is essential for neuronal function and maintenance, which becomes a challenge in these long axons. Membrane trafficking is responsible for a correct distribution and degradation of proteins, lipids and complex carbohydrates, and relies on membrane bound vesicles. The proper control and degradation of these vesicles, trafficking between pre- or post-synaptic compartments, is essential because failure to do so could lead to the accumulation of

dysfunctional cargoes or intracellular machinery, which may be toxic to the neuron. As neurons are long-lived post mitotic cells, that cannot transfer dysfunctional proteins and damaged organelles to daughter cells, seem more sensitive to their accumulation (Schwarz and Patrick, 2012; Wang et al., 2013). The proper functioning of the ER that is responsible, among other roles, for the synthesis, quality control and trafficking of proteins, is also essential for the neurons as we can see by the large number of neurodegenerative diseases associated with disturbances in ER dynamics (Yamanaka and Nukina, 2018) and stress response (Lindholm et al., 2006; Xiang et al., 2017). There are also a variety of examples in which ER and another organelle, mitochondria, are affected (Lindholm et al., 2006; Liu and Zhu, 2017). Mitochondria is another organelle very critical for the proper functioning of the neuron considering their intense energy demand. Dysfunction of mitochondrial dynamics and of their stress response pathway are implicated in several neurodegenerative diseases (Johri and Beal, 2012; Jovaisaite et al., 2014). All the above-mentioned processes are dependent on lipid homeostasis. Indeed, lipids are essential for most neuronal functions, development, and plasticity. Their abnormal metabolism is linked to several neurodegenerative disorders, including Alzheimer's and Parkinson's disease (Xu and Huang, 2006). The role of myelin, a membrane of the oligodendrocytes full of lipids, that facilitate a proper neuronal connectivity seem to be also important for axonal survival as suggested by demyelinating diseases such as Charcot-Marie-Tooth type 1 (CMT1) and multiple sclerosis (Ettle et al., 2016; Ravera and Panfoli, 2015). DNA repair and nucleotide metabolism, although only recently associated with HSP, seem to also be a common link to neurodegeneration. Although the molecular mechanisms underlying the pathogenicity is still unknown for the nucleotide metabolism, the explanation could be related to cellular and mitochondrial damage (Micheli et al., 2011; Ross and Truant, 2017).

The identification of new genes could potentially provide additional insights to some of these pathways or even identify novel pathways associated with hereditary spastic paraplegia. A challenge remains to find the commonalities in the pathological mechanisms in order to explain the similarities at the clinical level. This may facilitate also the discovery of common therapies for multiple genetic HSP entities, which are still lacking at the moment.

5.3.2 New spastic paraplegia genes and mechanisms

Our study resulted in the association for the first time, of *AIMP1* variants with hereditary spastic paraplegia. In addition, we were able to identify one family with *ALDH18A1* variants, also related for the first time with this disorder.

*AIMP1 (*Aminoacyl tRNA synthetase complex interacting multifunctional protein 1)/p43 encodes a multi-functional polypeptide with cytokine and tRNA-binding activities (Quevillon et

al., 1997). AIMP1 can be secreted as a cytokine to control angiogenesis, inflammation and wound healing, and as a glucagon-like hormone for glucose homeostasis (Park et al., 2006). It is also a component of the multi-tRNA synthetase complex that catalyzes the ligation of amino acids to their associated tRNAs. These aminoacyl tRNA synthetases have been associated with some neuronal diseases such as CMT and also to some autoimmune diseases and cancer (Park et al., 2008). However, for HSP the most interesting fact was its suggested role in maintaining axonal integrity. AIMP1 is expressed in the neurons of the brain and spinal cord, where it interacts with the neurofilament-light subunit seeming to be capable of regulating its phosphorylation. Neurofilaments (NFs) are part of the main cytoskeletal network maintaining the structural integrity of neurons that have critical roles in neuronal morphogenesis, neuronal migration, axonal growth, synaptic plasticity, and intracellular transport. Alteration of the phosphorylation level of the NFs by *AIMP1* resulted in disassembly of the NF network. So, by influencing NF assembly, *AIMP1* seems to regulate axon development and maintenance, which explains its dysfunction leading to axon degeneration associated with motor neuron degenerative diseases as suggested by the mice phenotype (Zhu et al., 2009).

ALDH18A1 (Aldehyde Dehydrogenase 18 Family Member A1) encodes an enzyme (delta-1-pyrroline-5-carboxylate synthase - P5CS) that catalyzes the first and critical step of proline and ornithine biosynthesis from glutamate. Interestingly, variants in several genes encoding enzymes of this pathway were associated with neurodegenerative disorders. This includes *PYCR1* associated with a neurocutaneous syndrome (Reversade et al., 2009) and *PYCR2* (Pyrroline-5-Carboxylate Reductase 2), identified as a candidate gene in one of our families, that was previously associated with a hypomyelinating leukodystrophy (Nakayama et al., 2015), being both genes associated with the last step of proline biosynthesis. Additionally, variants in the *ARG1* gene, that catalyzes the hydrolysis of arginine to ornithine and urea, were already identified in patients presenting a phenotype that includes spastic paraparesis (Baranello et al., 2014). Although the association of this P5CS pathways, as well as the urea cycle, to the HSP pathogenesis is not fully understood, it seems that the proline metabolism is vital for the adaptation to osmotic and cellular redox stress and seems to be involved in the apoptosis pathway (Fichman et al., 2015; Hu et al., 2008a, 2008b; Yasuda et al., 2013). A role in the mitochondria could also be hypothesized due to its localization in the mitochondrial inner membrane and to the findings, in *PYCR1* affected fibroblasts, of alterations in mitochondrial morphology and decreased membrane potential, together with increased rate of apoptosis upon oxidative stress (Reversade et al., 2009). *PYCR2* loss of function was also associated with a decreased mitochondrial membrane potential and increased susceptibility to apoptosis upon oxidative stress (Nakayama et al., 2015). However, up to now only mitochondria abnormalities were reported by Fisher et al. (Fischer et al., 2014) that showed swollen

mitochondria detected by electron microscopy in skin biopsies fibroblasts that could not be observed when cultured.

The *UCHL1* (Ubiquitin C-Terminal Hydrolase L1) gene, although only considered as a candidate gene in one of our families, was recently linked to HSP (SPG79) (Rydning et al., 2017). This gene encodes an enzyme highly expressed in neurons (Wilkinson et al., 1989), which has both hydrolase activity and ubiquitin ligase activity (Liu et al., 2002). This gene reinforces the importance of the ubiquitin-proteasome system and protein degradation in neurodegenerative disorders such as HSP.

Candidate genes

The remaining genes identified in the exome data are, up to now, only candidate genes for HSP but that constitute, in some cases, strong new candidates that could bring new insights into the disease pathogenicity, which is the case of *AKAP6*, *SUCO*, *SGSM3*, *MAP1A,* and *PYCR2,* already mentioned.

AKAP6 (A-Kinase Anchoring Protein 6) encodes the muscle A-kinase anchoring protein (mAKAP), a scaffolding protein known to interact with a large number of signaling molecules such as cAMP-dependent protein kinase (PKA), cAMP-specific phosphodiesterase type 4D3 (PDE4D3), 3-phosphoinositide-dependent protein kinase-1 (PDK1), among others. The interaction between AKAP6 and the cAMP-signaling pathway that is involved in modulating protein degradation via the ubiquitin/proteasome pathway (Huang et al., 2013), could possibly explain its role in neuroprotection.

SUCO (SUN Domain Containing Ossification Factor) is also an interesting candidate gene where a role in neuronal development was recently suggested (Sha et al., 2015). Although the pathogenic mechanism of *SUCO* variants is not known, it could be related with its subcellular localization in the rough endoplasmic reticulum. The ER is an important organelle already related to the pathophysiological mechanism of HSPs, by several genes such as *ATL1, SPAST, RTN2, REEP1* and *REEP2* involved in shaping the ER network, or genes *TFG, ARL6IP1, RAB3GAP2, BSCL2, ERLIN1* and *ERLIN2* involved in other components of the ER morphogenesis. So, a potential destabilization of this organelle could be the basis of the *SUCO* pathogenesis.

SGSM3 (Small G Protein Signaling Modulator 3) encodes a ubiquitous protein that, together with *SGSM1* and *SGSM2*, are expressed in the CNS neurons and suggested modulators of the small G protein RAP and RAB, which are known to mediate neuronal signal transduction and vesicular transportation pathways (Yang et al., 2007). This seems to indicate a role of this protein in membrane trafficking, mostly if we consider it could be modeling members of the Rab small GTPase family, which are needed to provide specificity to the intracellular destination of vesicles.

MAP1A (Microtubule Associated Protein 1A) is also an interesting candidate gene due to its expected role in the axonal transport. This gene encodes a microtubule- associated protein, which are known to be important for the microtubules stability, that is highly expressed in the brain (Fukuyama and Rapoport, 1995; Oudega et al., 1998). The presence of a homozygous *MAP1A* variant in mice resulted in a phenotype with tremors, ataxia, and loss of cerebellar Purkinje neurons in aged mice. This study also demonstrated the importance of this protein for the maintenance of the neuronal microtubule network (Liu et al., 2015).

In the remaining candidates, some could also potentially be part of some pathways already related with HSP, which is the case of *IFT74, BMP6, SERINC4, NPB* and *ABCA12* genes. The *IFT74* (Intraflagellar Transport 74) gene could have a potential role in microtubule dynamics as, together with *IFT81,* seem to be necessary for the tubulin transport within cilia (Bhogaraju et al., 2013). *BMP6* (Bone morphogenetic protein 6) encodes a member of the transforming growth factor-beta (TGF-β) super family, known to promote osteogenesis (Cheng et al., 2003; Ebisawa et al., 1999). This is an interesting candidate due to the fact that the BMP signaling has a critical role in several steps of brain development, with alterations in this pathway implicated in several neurodegenerative disorders, including HSP (Bayat et al., 2011; Liu and Niswander, 2005; Osses and Henríquez, 2014). The inhibition of this pathway by proteins encoded by genes like *NIPA1, ATL1, SPAST* and *SPG20* seems to be important for axon architecture and stability (Fassier et al., 2010; Tsang et al., 2009). A role for this protein in iron metabolism (Andriopoulos et al., 2009) and glucose homeostasis (Paralkar et al., 2010) was also suggested. *SERINC4* (Serine Incorporator 4) could have a potential role on lipid metabolism due to its part in the incorporation of serine into membranes and in facilitating the synthesis of phosphatidylserine and sphingolipids (Inuzuka et al., 2005). *NPB* (neuropeptide B) and *ABCA12* (ATP-Binding Cassette Subfamily A Member 12) could potentially have a role in membrane trafficking due to their protein functions as a neurotransmitter and an ATP-binding cassette transporter, respectively.

The identification of all these new genes and candidate genes associated with HSP is relevant not only because it leads to a better knowledge of the disease but also as it could potentially lead to the identification of new therapeutic targets, critical for the development of new treatments. The BMP signaling pathway is one interesting example as it appears to be a good target, amenable to modulation by small molecules (Blackstone, 2018).

6. Conclusions

6.1 Final remarks

Hereditary spastic paraplegias are a group of rare neurodegenerative disorders clinically and genetically very heterogeneous. Although rare, at least 193 families were identified with the disease in Portugal, 159 of which with samples available for genetic study. A screening of the most common genetic causes of HSP resulted in the molecular diagnosis of 59 families prior to this thesis.

This work, focused on the 100 Portuguese HSP families still without diagnosis, allowed us to:

1- Provide a molecular diagnosis to 24 families with HSP, which has a major impact on these patients and families that can now have a more specific follow-up, mostly in terms of genetic counseling. Also, it could improve their quality of life as it has been shown that, living with a non-identified condition is an additional source of stress.
2- Associate, for the first time, *AIMP1* variants with hereditary spastic paraplegia, which may allow the establishing of the molecular diagnosis of additional families worldwide.
3- Identify new candidate genes for HSP in 7 families, two of which (*AKAP6* and *SUCO*), we are further studying in order to ascertain their role in the HSP pathophysiology.
4- Detect new causing variants in genes already described in the literature, which increases the knowledge on the gene association with HSP.
5- Identify variants of unknown significance in 41 of the 69 families without molecular diagnosis (that did not proceed for exome study), that could potentially explain or partially explain the phenotype of some of these families after further study.
6- Suggest for the first time a new mode of inheritance for *KIF5A* variants, which, after functional study, could provide a molecular diagnosis to at least one additional family.
7- Further confirm the high genetic heterogeneity present in the Portuguese cohort, that excluding the large frequencies of variants in *SPG4* and *SPG11*, are distributed among at least 18 different HSP-causing genes, a number that will grow as a large part of the Portuguese HSP families are still without molecular diagnosis.

These accomplishments allowed us to fulfill this project major goals. Though, as almost half of the families remain without a definitive molecular diagnosis after the large screening performed, it shows that there is still a large path to pursuit. This is a huge task that could only be achieved with the use of large screening strategies like exome sequencing, but which are also not 100% effective. This study also highlights the importance of identifying new genes in HSP and of further studying genes already identified, in order to increase the knowledge on these gene functions and mechanisms, improving the variant reclassification and reducing the variants of unknown significance.

In conclusion, this work contributed to increase the knowledge on the genetics of HSP with the identification of additional genes associated with the disease, and strong candidates to be further studied. The study of these new genes and candidates is also relevant as it could contribute to a better understanding of the pathophysiological mechanisms involved in HSP and to the identification of new therapeutic targets.

6.2 Future perspectives

This work resulted in the identification of several interesting candidate genes and variants of unknown significance in the Portuguese families that need to be further studied. However, as it is almost impossible to explore all of them, we select the stronger ones for further analysis. Our main focus will now be to functionally characterize the two main candidate genes identified, *AKAP6* and *SUCO*.

Fibroblasts from these two families were collected and are now available to study. As for the *AKAP6* gene, the use of patient fibroblasts as a cellular model has proven not to be an effective strategy; thus, we are now exploring the possibility of generating induced pluripotent stem cells from the patients' fibroblasts and differentiate them into neuronal cells. The functional study of the patient cells from these two families could reveal cellular abnormalities, allowing the establishment of these variants functional consequences and the potential role for these genes in HSP pathogenesis.

If additional proof for *SUCO* pathogenicity is found, we will screen the remaining Portuguese cohort still without molecular diagnosis for this gene, to identify additional families whose phenotype could be explained by variants in this gene.

It is also one of our future aims to characterize the homozygous variant identified in *KIF5A* (family S88), in order to ascertain the possibility of a new inheritance mode for variants in this gene.

References

Abou Jamra, R., Philippe, O., Raas-Rothschild, A., Eck, S.H., Graf, E., Buchert, R., Borck, G., Ekici, A., Brockschmidt, F.F., Nöthen, M.M., et al. (2011). Adaptor protein complex 4 deficiency causes severe autosomal-recessive intellectual disability, progressive spastic paraplegia, shy character, and short stature. Am. J. Hum. Genet. *88*, 788–795.

Akiyama, M. (2010). ABCA12 mutations and autosomal recessive congenital ichthyosis: A review of genotype/phenotype correlations and of pathogenetic conceptsa. Hum. Mutat. *31*, 1090–1096.

Akizu, N., Cantagrel, V., Schroth, J., Cai, N., Vaux, K., McCloskey, D., Naviaux, R.K., Van Vleet, J., Fenstermaker, A.G., Silhavy, J.L., et al. (2013). AMPD2 Regulates GTP Synthesis and Is Mutated in a Potentially Treatable Neurodegenerative Brainstem Disorder. Cell *154*, 505–517.

Alazami, A.M., Adly, N., Al Dhalaan, H., and Alkuraya, F.S. (2011). A nullimorphic ERLIN2 mutation defines a complicated hereditary spastic paraplegia locus (SPG18). Neurogenetics *12*, 333–336.

Aldahmesh, M.A., Mohamed, J.Y., Alkuraya, H.S., Verma, I.C., Puri, R.D., Alaiya, A.A., Rizzo, W.B., and Alkuraya, F.S. (2011). Recessive mutations in ELOVL4 cause ichthyosis, intellectual disability, and spastic quadriplegia. Am. J. Hum. Genet. *89*, 745–750.

Alderson, N.L., Maldonado, E.N., Kern, M.J., Bhat, N.R., and Hama, H. (2006). FA2H-dependent fatty acid 2-hydroxylation in postnatal mouse brain. J. Lipid Res. *47*, 2772–2780.

Andriopoulos, B., Corradini, E., Xia, Y., Faasse, S.A., Chen, S., Grgurevic, L., Knutson, M.D., Pietrangelo, A., Vukicevic, S., Lin, H.Y., et al. (2009). BMP6 is a key endogenous regulator of hepcidin expression and iron metabolism. Nat. Genet. *41*, 482–487.

Annilo, T., Shulenin, S., Chen, Z.Q., Arnould, I., Prades, C., Lemoine, C., Maintoux-Larois, C., Devaud, C., Dean, M., Denèfle, P., et al. (2002). Identification and characterization of a novel ABCA subfamily member, ABCA12, located in the lamellar ichthyosis region on 2q34. Cytogenet. Genome Res. *98*, 169–176.

Arif, B., Kumar, K.R., Seibler, P., Vulinovic, F., Fatima, A., Winkler, S., Nürnberg, G., Thiele, H., Nürnberg, P., Jamil, A.Z., et al. (2013). A Novel OPA3 Mutation Revealed by Exome Sequencing. JAMA Neurol. *70*, 783.

Armstrong, L., Biancheri, R., Shyr, C., Rossi, A., Sinclair, G., Ross, C.J., Tarailo-Graovac, M., Wasserman, W.W., and Van Karnebeek, C.D.M. (2014). AIMP1 deficiency presents as a cortical neurodegenerative disease with infantile onset. Neurogenetics *15*, 157–159.

Atorino, L., Silvestri, L., Koppen, M., Cassina, L., Ballabio, A., Marconi, R., Langer, T., and Casari, G. (2003). Loss of m-AAA protease in mitochondria causes complex I deficiency and increased sensitivity to oxidative stress in hereditary spastic paraplegia. J. Cell Biol. *1177711*,

21–9525.

Baas, P.W., Nadar, C.V., and Myers, K.A. (2006). Axonal transport of microtubules: The long and short of it. Traffic *7*, 490–498.

Babin, P.J., Goizet, C., and Raldú, D. (2014). Zebrafish models of human motor neuron diseases: Advantages and limitations. Prog. Neurobiol. *118*, 36–58.

Bache, K.G., Slagsvold, T., Cabezas, A., Rosendal, K.R., Raiborg, C., and Stenmark, H. (2004). The growth-regulatory protein HCRP1/hVps37A is a subunit of mammalian ESCRT-I and mediates receptor down-regulation. Mol. Biol. Cell *15*, 4337–4346.

Baranello, G., Alfei, E., Martinelli, D., Rizzetto, M., Cazzaniga, F., Dionisi-Vici, C., Gellera, C., and Castellotti, B. (2014). Hyperargininemia: 7-month follow-up under sodium benzoate therapy in an Italian child presenting progressive spastic paraparesis, cognitive decline, and novel mutation in arg1 gene. Pediatr. Neurol. *51*, 430–433.

Bathgate, D., Yu-Wai-Man, P., Webb, B., Taylor, R.W., Fowler, B., and Chinnery, P.F. (2012). Recessive spastic paraparesis associated with complex I deficiency due to MTHFR mutations. J. Neurol. Neurosurg. Psychiatry *83*, 115.

Bayat, V., Jaiswal, M., and Bellen, H.J. (2011). The BMP signaling pathway at the Drosophila neuromuscular junction and its links to neurodegenerative diseases. Curr. Opin. Neurobiol. *21*, 182–188.

Bayrakli, F., Poyrazoglu, H.G., Yuksel, S., Yakicier, C., Erguner, B., Sagiroglu, M.S., Yuceturk, B., Ozer, B., Doganay, S., Tanrikulu, B., et al. (2015). Hereditary spastic paraplegia with recessive trait caused by mutation in KLC4 gene. J. Hum. Genet. *60*, 763–768.

Beetz, C., Johnson, A., Schuh, A.L., Thakur, S., Varga, R.-E., Fothergill, T., Hertel, N., Bomba-Warczak, E., Thiele, H., Nürnberg, G., et al. (2013a). Inhibition of TFG function causes hereditary axon degeneration by impairing endoplasmic reticulum structure. Proc. Natl. Acad. Sci. U. S. A. *110*, 5091–5096.

Beetz, C., Koch, N., Khundadze, M., Zimmer, G., Nietzsche, S., Hertel, N., Huebner, A.-K., Mumtaz, R., Schweizer, M., Dirren, E., et al. (2013b). A spastic paraplegia mouse model reveals REEP1-dependent ER shaping. J. Clin. Invest. *123*, 4273–4282.

Bento, C.F., Ashkenazi, A., Jimenez-Sanchez, M., and Rubinsztein, D.C. (2016). The Parkinson's disease-associated genes ATP13A2 and SYT11 regulate autophagy via a common pathway. Nat. Commun. *7*, 11803.

Bhogaraju, S., Cajanek, L., Fort, C., Blisnick, T., Weber, K., Taschner, M., Mizuno, N., Lamla, S., Bastin, P., Nigg, E.A., et al. (2013). Molecular basis of tubulin transport within the cilium by IFT74 and IFT81. Science (80-.). *341*, 1009–1012.

Biancheri, R., Rossi, A., Zara, F., and Filocamo, M. (2011). AIMP1/p43 mutation and PMLD. Am. J. Hum. Genet. *88*, 391.

Bilguvar, K., Tyagi, N.K., Ozkara, C., Tuysuz, B., Bakircioglu, M., Choi, M., Delil, S., Caglayan,

A.O., Baranoski, J.F., Erturk, O., et al. (2013). Recessive loss of function of the neuronal ubiquitin hydrolase UCHL1 leads to early-onset progressive neurodegeneration. Proc. Natl. Acad. Sci. U. S. A. *110*, 3489–3494.

Bitoun, M., Durieux, A.-C., Prudhon, B., Bevilacqua, J.A., Herledan, A., Sakanyan, V., Urtizberea, A., Cartier, L., Romero, N.B., and Guicheney, P. (2009). Dynamin 2 mutations associated with human diseases impair clathrin-mediated receptor endocytosis. Hum. Mutat. *30*, 1419–1427.

Blackstone, C. (2012). Cellular pathways of hereditary spastic paraplegia. Annu. Rev. Neurosci. *35*, 25–47.

Blackstone, C. (2018). Hereditary spastic paraplegia. Handb. Clin. Neurol. *148*, 633–652.

Blackstone, C., Kane, C.J.O., Reid, E., O'Kane, C.J., and Reid, E. (2011). Hereditary spastic paraplegias: membrane traffic and the motor pathway. Nat. Rev. Neurosci. *12*, 31–42.

Boespflug-Tanguy, O., Aubourg, P., Dorboz, I., Bégou, M., Giraud, G., Sarret, C., and Vaurs-Barrière, C. (2011). Neurodegenerative disorder related to AIMP1/p43 mutation is not a PMLD. Am. J. Hum. Genet. *88*, 392–393.

Bonifacino, J.S., and Glick, B.S. (2004). The Mechanisms of Vesicle Budding and Fusion. Cell *116*, 153–166.

Bot, S.T. de, Schelhaas, H.J., Kamsteeg, E.J., and Van De Warrenburg, B.P.C. (2012). Hereditary spastic paraplegia caused by a mutation in the VCP gene. Brain *135*, 1–3.

Botzolakis, E.J., Zhao, J., Gurba, K.N., Macdonald, R.L., and Hedera, P. (2012). The effect of HSP-causing mutations in SPG3A and NIPA1 on the assembly, trafficking, and interaction between atlastin-1 and NIPA1. Mol. Cell. Neurosci. *46*, 122–135.

Bouhlal, Y., Amouri, R., El Euch-Fayeche, G., Hentati, F., Bosgoed, E.A., Cremers, F.P., and al., et (2011). Autosomal recessive spastic ataxia of Charlevoix-Saguenay: an overview. Parkinsonism Relat. Disord. *17*, 418–422.

Bouhouche, a, Benomar, a, Bouslam, N., Chkili, T., and Yahyaoui, M. (2006). Mutation in the epsilon subunit of the cytosolic chaperonin-containing t-complex peptide-1 (Cct5) gene causes autosomal recessive mutilating sensory neuropathy with spastic paraplegia. J. Med. Genet. *43*, 441–443.

Boukhris, A., Schule, R., Loureiro, J.L., Lourenço, C.M., Mundwiller, E., Gonzalez, M.A., Charles, P., Gauthier, J., Rekik, I., Acosta Lebrigio, R.F., et al. (2013). Alteration of Ganglioside Biosynthesis Responsible for Complex Hereditary Spastic Paraplegia.

Branchu, J., Boutry, M., Sourd, L., Depp, M., Leone, C., Corriger, A., Vallucci, M., Esteves, T., Matusiak, R., Dumont, M., et al. (2017). Loss of spatacsin function alters lysosomal lipid clearance leading to upper and lower motor neuron degeneration. Neurobiol. Dis. *102*, 21–37.

Bross, P., Naundrup, S., Hansen, J., Nielsen, M.N., Christensen, J.H., Kruhøffer, M., Palmfeldt, J., Corydon, T.J., Gregersen, N., Ang, D., et al. (2008). The Hsp60-(p.V98I) mutation

associated with hereditary spastic paraplegia SPG13 compromises chaperonin function both in vitro and in vivo. J. Biol. Chem. *283*, 15694–15700.

Browman, D.T., Resek, M.E., Zajchowski, L.D., and Robbins, S.M. (2006). Erlin-1 and erlin-2 are novel members of the prohibitin family of proteins that define lipid-raft-like domains of the ER. J. Cell Sci. *119*, 3149–3160.

Carpenter, A.E., Jones, T.R., Lamprecht, M.R., Clarke, C., Kang, I.H., Friman, O., Guertin, D. a, Chang, J.H., Lindquist, R. a, Moffat, J., et al. (2006). CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biol. *7*, R100.

Casari, G., De Fusco, M., Ciarmatori, S., Zeviani, M., Mora, M., Fernandez, P., De Michele, G., Filla, A., Cocozza, S., Marconi, R., et al. (1998). Spastic paraplegia and OXPHOS impairment caused by mutations in paraplegin, a nuclear-encoded mitochondrial metalloprotease. Cell *93*, 973–983.

Chang, P.A., and Wu, Y.J. (2010). Neuropathy target esterase: An essential enzyme for neural development and axonal maintenance. Int. J. Biochem. Cell Biol. *42*, 573–575.

Chang, J., Lee, S., and Blackstone, C. (2014). Spastic paraplegia proteins spastizin and spatacsin mediate autophagic lysosome reformation. J. Clin. Invest. *124*, 5249–5262.

Chen, M., and Xu, R. (2011). Motor coordination deficits in Alpk1 mutant mice with the inserted piggyBac transposon. BMC Neurosci. *12*, 1.

Cheng, H., Jiang, W., Phillips, F.M., Haydon, R.C., Peng, Y., Zhou, L., Luu, H.H., An, N., Breyer, B., Vanichakarn, P., et al. (2003). Osteogenic activity of the fourteen types of human bone morphogenetic proteins (BMPs). J. Bone Joint Surg. Am. *85–A*, 1544–1552.

Cheon, C.K., Lim, S.-H., Kim, Y.-M., Kim, D., Lee, N.-Y., Yoon, T.-S., Kim, N.-S., Kim, E., and Lee, J.-R. (2017). Autosomal dominant transmission of complicated hereditary spastic paraplegia due to a dominant negative mutation of KIF1A, SPG30 gene. Sci. Rep. *7*, 12527.

Chrestian, N., Dupré, N., Gan-Or, Z., Szuto, A., Chen, S., Venkitachalam, A., Brisson, J.-D., Warman-Chardon, J., Ahmed, S., Ashtiani, S., et al. (2017). Clinical and genetic study of hereditary spastic paraplegia in Canada. Neurol. Genet. *3*, e122.

Ciccarelli, F.D., Proukakis, C., Patel, H., Cross, H., Azam, S., Patton, M.A., Bork, P., and Crosby, A.H. (2003). The identification of a conserved domain in both spartin and spastin, mutated in hereditary spastic paraplegia. Genomics *81*, 437–441.

Citterio, A., Arnoldi, A., Panzeri, E., Merlini, L., D'Angelo, M.G., Musumeci, O., Toscano, A., Bondi, A., Martinuzzi, A., Bresolin, N., et al. (2015). Variants in KIF1A gene in dominant and sporadic forms of hereditary spastic paraparesis. J. Neurol. *262*, 2684–2690.

Clemen, C.S., Tangavelou, K., Strucksberg, K.-H., Just, S., Gaertner, L., Regus-Leidig, H., Stumpf, M., Reimann, J., Coras, R., Morgan, R.O., et al. (2010). Strumpellin is a novel valosin-containing protein binding partner linking hereditary spastic paraplegia to protein aggregation diseases. Brain *133*, 2920–2941.

Corona, P., Lamantea, E., Greco, M., Carrara, F., Agostino, A., Guidetti, D., Dotti, M.T., Mariotti, C., and Zeviani, M. (2002). Novel heteroplasmic mtDNA mutation in a family with heterogeneous clinical presentations. Ann. Neurol. *51*, 118–122.

Coutelier, M., Goizet, C., Durr, A., Habarou, F., Morais, S., Dionne-Laporte, A., Tao, F., Konop, J., Stoll, M., Charles, P., et al. (2015). Alteration of ornithine metabolism leads to dominant and recessive hereditary spastic paraplegia. Brain *138*, 2191–2205.

Coutelier, M., Hammer, M.B., Stevanin, G., Monin, M.-L., Davoine, C.-S., Mochel, F., Labauge, P., Ewenczyk, C., Ding, J., Gibbs, J.R., et al. (2018). Efficacy of Exome-Targeted Capture Sequencing to Detect Mutations in Known Cerebellar Ataxia Genes. JAMA Neurol.

Coutinho, P., Barros, J., Zemmouri, R., Guimarães, J., Alves, C., Chorão, R., Lourenço, E., Ribeiro, P., Loureiro, J.L., Santos, J. V., et al. (1999). Clinical Heterogeneity of Autosomal Recessive Spastic Paraplegias. Arch. Neurol. *56*, 943–949.

Coutinho, P., Ruano, L., Loureiro, J.L., Cruz, V.T., Barros, J., Tuna, A., Barbot, C., Guimarães, J., Alonso, I., Silveira, I., et al. (2013). Hereditary ataxia and spastic paraplegia in Portugal: a population-based prevalence study. JAMA Neurol. *70*, 746–755.

Crosby, A.H., and Proukakis, C. (2002). Is the transportation highway the right road for hereditary spastic paraplegia? Am. J. Hum. Genet. *71*, 1009–1016.

Crow, Y.J., Zaki, M.S., Ghada, M.S.A., Boesp, O., Cordeiro, N.J. V, Gleeson, J.G., Rani, N., Laugel, V., Renaldo, F., Rodriguez, D., et al. (2014). Mutations in ADAR1 , IFIH1 , and RNASEH2B Presenting As Spastic Paraplegia Crow et al . Mutations in ADAR1 , IFIH1 , and RNASEH2B Presenting As Spastic Paraplegia Crow et al . Neuropediatrics *45*, 386–393.

Dahme, M., Bartsch, U., Martini, R., Anliker, B., Schachner, M., and Mantei, N. (1997). Disruption of the mouse L1 gene leads to malformations of the nervous system. Nat. Genet. *17*, 346–349.

Dale, R.C., Gornall, H., Singh-Grewal, D., Alcausin, M., Rice, G.I., and Crow, Y.J. (2010). Familial Aicardi-Goutières syndrome due to SAMHD1 mutations is associated with chronic arthropathy and contractures. Am. J. Med. Genet. Part A *152*, 938–942.

DeLuca, G.C., Ebers, G.C., and Esiri, M.M. (2004). The extent of axonal loss in the long tracts in hereditary spastic paraplegia. Neuropathol. Appl. Neurobiol. *30*, 576–584.

Denora, P.S., Smets, K., Zolfanelli, F., Ceuterick-de Groote, C., Casali, C., Deconinck, T., Sieben, A., Gonzales, M., Zuchner, S., Darios, F., et al. (2016). Motor neuron degeneration in spastic paraplegia 11 mimics amyotrophic lateral sclerosis lesions (Oxford University Press).

Denton, K.R., Xu, C., Shah, H., and Li, X.-J. (2016). Modeling Axonal Defects in Hereditary Spastic Paraplegia with Human Pluripotent Stem Cells. Front. Biol. (Beijing). *11*, 339–354.

Depienne, C., Tallaksen, C., Lephay, J.Y., Bricka, B., Poea-Guyon, S., Fontaine, B., Labauge, P., Brice, A., and Durr, A. (2006). Spastin mutations are frequent in sporadic spastic paraparesis and their spectrum is different from that observed in familial cases. J. Med. Genet.

43, 259–265.

Depienne, C., Stevanin, G., Brice, A., and Durr, A. (2007). Hereditary spastic paraplegias: an update. Curr. Opin. Neurol. *20*, 674–680.

Dion, P.A., Daoud, H., and Rouleau, G.A. (2009). Genetics of motor neuron disorders: new insights into pathogenic mechanisms. Nat. Rev. Genet. *10*, 769–782.

Dodge, K.L., Khouangsathiene, S., Kapiloff, M.S., Mouton, R., Hill, E. V, Houslay, M.D., Langeberg, L.K., and Scott, J.D. (2001). mAKAP assembles a protein kinase A/PDE4 phosphodiesterase cAMP signaling module. EMBO J. *20*, 1921–1930.

Dodge-Kafka, K.L., Soughayer, J., Pare, G.C., Carlisle Michel, J.J., Langeberg, L.K., Kapiloff, M.S., and Scott, J.D. (2005). The protein kinase A anchoring protein mAKAP coordinates two integrated cAMP effector pathways. Nature *437*, 574–578.

Dong, C., Wei, P., Jian, X., Gibbs, R., Boerwinkle, E., Wang, K., and Liu, X. (2015). Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies. Hum. Mol. Genet. *24*, 2125–2137.

Duncan, I.D., Bugiani, M., Radcliff, A.B., Moran, J.J., Lopez-Anido, C., Duong, P., August, B.K., Wolf, N.I., van der Knaap, M.S., and Svaren, J. (2017a). A mutation in the *Tubb4a* gene leads to microtubule accumulation with hypomyelination and demyelination. Ann. Neurol. *81*, 690–702.

Duncan, I.D., Bugiani, M., Radcliff, A.B., Moran, J.J., Lopez-anido, C., Duong, P., August, B.K., Wolf, N.I., van der Knaap, M.S., Svaren, J., et al. (2017b). A mutation in the Tubb4a gene leads to microtubule accumulation with hypomyelination and demyelination. Ann. Neurol. *81*, 1–43.

Durand, C.M., Dhers, L., Tesson, C., Tessa, A., Fouillen, L., Jacqueré, S., Raymond, L., Coupry, I., Benard, G., Darios, F., et al. (2018). CYP2U1 activity is altered by missense mutations in hereditary spastic paraplegia 56. Hum. Mutat. *39*, 140–151.

Eastman, S.W., Yassaee, M., and Bieniasz, P.D. (2009). A role for ubiquitin ligases and Spartin/SPG20 in lipid droplet turnover. J. Cell Biol. *184*, 881–894.

Ebbing, B., Mann, K., Starosta, A., Jaud, J., Schöls, L., Schüle, R., and Woehlke, G. (2008). Effect of spastic paraplegia mutations in KIF5A kinesin on transport activity. Hum. Mol. Genet. *17*, 1245–1252.

Ebisawa, T., Tada, K., Kitajima, I., Tojo, K., Sampath, T.K., Kawabata, M., Miyazono, K., and Imamura, T. (1999). Characterization of bone morphogenetic protein-6 signaling pathways in osteoblast differentiation. J. Cell Sci. *112*.

Edvardson, S., Hama, H., Shaag, A., Gomori, J.M., Berger, I., Soffer, D., Korman, S.H., Taustein, I., Saada, A., and Elpeleg, O. (2008). Mutations in the Fatty Acid 2-Hydroxylase Gene Are Associated with Leukodystrophy with Spastic Paraparesis and Dystonia. Am. J. Hum. Genet. *83*, 643–648.

Elleuch, N., Depienne, C., Benomar, A., Hernandez, A.M.O., Ferrer, X., Fontaine, B., Grid, D., Tallaksen, C.M.E., Zemmouri, R., Stevanin, G., et al. (2006). Mutation analysis of the paraplegin gene (SPG7) in patients with hereditary spastic paraplegia. Neurology *66*, 654–659.

Elsaid, M.F., Khalid, I., Chalhoub, N., Elsotouhy, A., Mudehk, N. El, and Aleem, A.A. (2017). NT5C2 novel splicing variant expands the phenotypic spectrum of Spastic Paraplegia (SPG45): case report of a new member of thin corpus callosum SPG-Subgroup. BMC Med. Genet. 18–33.

Elsayed, L.E.O., Mohammed, I.N., Hamed, A.A.A., Elseed, M.A., Johnson, A., Mairey, M., Mohamed, H.E.S.A., Idris, M.N., Salih, M.A.M., El-sadig, S.M., et al. (2017). Hereditary spastic paraplegias: identification of a novel SPG57 variant affecting TFG oligomerization and description of HSP subtypes in Sudan. Eur. J. Hum. Genet. *25*, 100–110.

Erlich, Y., Edvardson, S., Hodges, E., Zenvirt, S., Thekkat, P., Shaag, A., Dor, T., Hannon, G.J., and Elpeleg, O. (2011). Exome sequencing and disease-network analysis of a single family implicate a mutation in KIF1A in hereditary spastic paraparesis. Genome Res. *21*, 658–664.

Errico, A., Ballabio, A., and Rugarli, E.I. (2002). Spastin, the protein mutated in autosomal dominant hereditary spastic paraplegia, is involved in microtubule dynamics. Hum. Mol. Genet. *11*, 153–163.

Esteves, T., Durr, A., Mundwiller, E., Loureiro, J.L., Boutry, M., Gonzalez, M.A., Gauthier, J., El-Hachimi, K.H., Depienne, C., Muriel, M.P., et al. (2014). Loss of association of REEP2 with membranes leads to hereditary spastic paraplegia. Am. J. Hum. Genet. *94*, 268–277.

Estrada-Cuzcano, A., Martin, S., Chamova, T., Synofzik, M., Timmann, D., Holemans, T., Andreeva, A., Reichbauer, J., De Rycke, R., Chang, D.-I., et al. (2017). Loss-of-function mutations in the *ATP13A2/* PARK9 gene cause complicated hereditary spastic paraplegia (SPG78). Brain *140*, 287–305.

Ettle, B., Schlachetzki, J.C.M., and Winkler, J. (2016). Oligodendroglia and Myelin in Neurodegenerative Diseases: More Than Just Bystanders? Mol. Neurobiol. *53*, 3046–3062.

Fan, Z., Greenwood, R., Felix, A.C.G., Shiloh-Malawsky, Y., Tennison, M., Roche, M., Crooks, K., Weck, K., Wilhelmsen, K., Berg, J., et al. (2014). GCH1 heterozygous mutation identified by whole-exome sequencing as a treatable condition in a patient presenting with progressive spastic paraplegia. J. Neurol. *261*, 622–624.

Fassier, C., Hutt, J.A., Scholpp, S., Lumsden, A., Giros, B., Nothias, F., Schneider-Maunoury, S., Houart, C., and Hazan, J. (2010). Zebrafish atlastin controls motility and spinal motor axon architecture via inhibition of the BMP pathway. Nat. Neurosci. *13*, 1380–1387.

Fasullo, M., and Endres, L. (2015). Nucleotide salvage deficiencies, DNA damage and neurodegeneration. Int. J. Mol. Sci. *16*, 9431–9449.

Feinstein, M., Markus, B., Noyman, I., Shalev, H., Flusser, H., Shelef, I., Liani-Leibson, K., Shorer, Z., Cohen, I., Khateeb, S., et al. (2010). Pelizaeus-merzbacher-like disease caused by AIMP1/p43 homozygous mutation. Am. J. Hum. Genet. *87*, 820–828.

Ferreirinha, F., Quattrini, A., Pirozzi, M., Valsecchi, V., Dina, G., Broccoli, V., Auricchio, A., Piemonte, F., Tozzi, G., Gaeta, L., et al. (2004). Axonal degeneration in paraplegin-deficient mice is associated with abnormal mitochondria and impairment of axonal transport. J. Clin. Invest. *113*, 231–242.

Fichman, Y., Gerdes, S.Y., Kovács, H., Szabados, L., Zilberstein, A., and Csonka, L.N. (2015). Evolution of proline biosynthesis: enzymology, bioinformatics, genetics, and transcriptional regulation. Biol. Rev. Camb. Philos. Soc. *90*, 1065–1099.

Filosto, M., Scarpelli, M., Cotelli, M.S., Vielmi, V., Todeschini, A., Gregorelli, V., Tonin, P., Tomelleri, G., and Padovani, A. (2011). The role of mitochondria in neurodegenerative diseases. J. Neurol. *258*, 1763–1774.

Fink, J.K. (2013). Hereditary spastic paraplegia: Clinico-pathologic features and emerging molecular mechanisms. Acta Neuropathol. *126*, 307–328.

Fischer, B., Callewaert, B., Schröter, P., Coucke, P.J., Schlack, C., Ott, C.-E., Morroni, M., Homann, W., Mundlos, S., Morava, E., et al. (2014). Severe congenital cutis laxa with cardiovascular manifestations due to homozygous deletions in ALDH18A1. Mol. Genet. Metab. *112*, 310–316.

Flanagan-Steet, H., Fox, M.A., Meyer, D., and Sanes, J.R. (2005). Neuromuscular synapses can form in vivo by incorporation of initially aneural postsynaptic specializations. Development *132*, 4471–4481.

Fukuyama, R., and Rapoport, S.I. (1995). Brain-specific expression of human microtubule-associated protein 1A (MAP1A) gene and its assignment to human chromosome 15. J. Neurosci. Res. *40*, 820–825.

Furuta, N., Ikeda, M., Hirayanagi, K., Fujita, Y., Amanuma, M., and Okamoto, K. (2012). A Novel GJA1 Mutation in Oculodentodigital Dysplasia with Progressive Spastic Paraplegia and Sensory Deficits. Intern. Med. *51*, 93–98.

Gan-Or, Z., Bouslam, N., Birouk, N., Lissouba, A., Chambers, D.B., Vérièpe, J., Androschuck, A., Laurent, S.B., Rochefort, D., Spiegelman, D., et al. (2016). Mutations in CAPN1 Cause Autosomal-Recessive Hereditary Spastic Paraplegia. Am. J. Hum. Genet. *98*, 1038–1046.

Gerondopoulos, A., Bastos, R.N., Yoshimura, S.I., Anderson, R., Carpanini, S., Aligianis, I., Handley, M.T., and Barr, F.A. (2014). Rab18 and a Rab18 GEF complex are required for normal ER structure. J. Cell Biol. *205*, 707–720.

Lo Giudice, T., Lombardi, F., Santorelli, F.M., Kawarai, T., and Orlacchio, A. (2014). Hereditary spastic paraplegia: Clinical-genetic characteristics and evolving molecular mechanisms. Exp. Neurol. *261*, 518–539.

Glória, V.G., Araújo, M.M. De, Santos, M., Leal, R., Almeida, S.F. De, Carmo, A.M., Moreira, A., Immunol, J., da Gloria, V.G., Martins de Araujo, M., et al. (2014). T Cell Activation Regulates CD6 Alternative Splicing by Transcription Dynamics and SRSF1. J. Immunol. *193*, 391–399.

Glynn, P. (2013). Neuronal phospholipid deacylation is essential for axonal and synaptic integrity. Biochim. Biophys. Acta - Mol. Cell Biol. Lipids *1831*, 633–641.

Goizet, C., Depienne, C., Benard, G., Boukhris, A., Mundwiller, E., Sol??, G., Coupry, I., Pilliod, J., Martin-N??grier, M.L., Fedirko, E., et al. (2011). REEP1 mutations in SPG31: Frequency, mutational spectrum, and potential association with mitochondrial morpho-functional dysfunction. Hum. Mutat. *32*, 1118–1127.

Gonzalez, M., Nampoothiri, S., Kornblum, C., Oteyza, A.C., Walter, J., Konidari, I., Hulme, W., Speziani, F., Schöls, L., Züchner, S., et al. (2013). Mutations in phospholipase DDHD2 cause autosomal recessive hereditary spastic paraplegia (SPG54). Eur. J. Hum. Genet. *21*, 1214–1218.

Goyal, U., and Blackstone, C. (2013). Untangling the web: Mechanisms underlying ER network formation. Biochim. Biophys. Acta - Mol. Cell Res. *1833*, 2492–2498.

Gregianin, E., Vazza, G., Scaramel, E., Boaretto, F., Vettori, A., Leonardi, E., Tosatto, S.C.E., Manara, R., Pegoraro, E., and Mostacciuolo, M.L. (2013). A novel SACS mutation results in non-ataxic spastic paraplegia and peripheral neuropathy. Eur. J. Neurol. *20*, 1486–1491.

Grünewald, A., Arns, B., Seibler, P., Rakovic, A., Münchau, A., Ramirez, A., Sue, C.M., and Klein, C. (2012). ATP13A2 mutations impair mitochondrial function in fibroblasts from patients with Kufor-Rakeb syndrome. Neurobiol. Aging *33*, 1843.e1-1843.e7.

Gudbjartsson, D.F., Jonasson, K., Frigge, M.L., and Kong, A. (2000). Allegro, a new computer program for multipoint linkage analysis. Nat. Genet. *25*, 12–13.

Gusdon, A.M., Zhu, J., Van Houten, B., and Chu, C.T. (2012). ATP13A2 regulates mitochondrial bioenergetics through macroautophagy. Neurobiol. Dis. *45*, 962–972.

Halevy, A., Lerer, I., Cohen, R., Kornreich, L., Shuper, A., Gamliel, M., Zimerman, B. El, Korabi, I., Meiner, V., Straussberg, R., et al. (2014). Novel EXOSC3 mutation causes complicated hereditary spastic paraplegia. J. Neurol. *261*, 2165–2169.

Hanein, S., Dürr, A., Ribai, P., Forlani, S., Leutenegger, A.L., Nelson, I., Babron, M.C., Elleuch, N., Depienne, C., Charon, C., et al. (2007). A novel locus for autosomal dominant “uncomplicated” hereditary spastic paraplegia maps to chromosome 8p21.1-q13.3. Hum. Genet. *122*, 261–273.

Hanein, S., Martin, E., Boukhris, A., Byrne, P., Goizet, C., Hamri, A., Benomar, A., Lossos, A., Denora, P., Fernandez, J., et al. (2008). Identification of the SPG15 Gene, Encoding Spastizin, as a Frequent Cause of Complicated Autosomal-Recessive Spastic Paraplegia, Including Kjellin Syndrome. Am. J. Hum. Genet. *82*, 992–1002.

Hansen, J., Corydon, T.J., Palmfeldt, J., Dürr, A., Fontaine, B., Nielsen, M.N., Christensen, J.H., Gregersen, N., and Bross, P. (2008). Decreased expression of the mitochondrial matrix proteases Lon and ClpP in cells from a patient with hereditary spastic paraplegia (SPG13). Neuroscience *153*, 474–482.

Hansen, J.J., Dürr, A., Cournu-Rebeix, I., Georgopoulos, C., Ang, D., Nielsen, M.N., Davoine, C.-S., Brice, A., Fontaine, B., Gregersen, N., et al. (2002). Hereditary Spastic Paraplegia SPG13 Is Associated with a Mutation in the Gene Encoding the Mitochondrial Chaperonin Hsp60. Am. J. Hum. Genet. *70*, 1328–1332.

Harbour, M.E., Breusegem, S.Y.A., Antrobus, R., Freeman, C., Reid, E., and Seaman, M.N.J. (2010). The cargo-selective retromer complex is a recruiting hub for protein complexes that regulate endosomal tubule dynamics. J. Cell Sci. *123*.

Harding, A.E. (1981). Hereditary “pure” spastic paraplegia: a clinical and genetic study of 22 families. J. Neurol. Neurosurg. Psychiatry *44*, 871–883.

Harding, A.E. (1983). Classification of the Hereditary Ataxias and Paraplegias. Lancet *321*, 1151–1155.

Harris, L.A.L.S., Skinner, J.R., and Wolins, N.E. (2013). Imaging of Neutral Lipids and Neutral Lipid Associated Proteins. Methods Cell Biol. *116*, 213–226.

Hazan, J., Fonknechten, N., Mavel, D., Paternotte, C., Samson, D., Artiguenave, F., Davoine, C.-S., Cruaud, C., Dürr, A., Wincker, P., et al. (1999). Spastin, a new AAA protein, is altered in the most frequent form of autosomal dominant spastic paraplegia. Nat. Genet. *23*, 296–303.

Heffernan, L.F., and Simpson, J.C. (2014). The trials and tubule-ations of Rab6 involvement in Golgi-to-ER retrograde transport. Biochem. Soc. Trans. *42*, 1453–1459.

Hensiek, A., Kirker, S., and Reid, E. (2015). Diagnosis, investigation and management of hereditary spastic paraplegias in the era of next-generation sequencing. J. Neurol. *262*, 1601–1612.

Herrmann, J.M., and Spang, A. (2015). Intracellular Parcel Service: Current Issues in Intracellular Membrane Trafficking. In Methods in Molecular Biology (Clifton, N.J.), pp. 1–12.

Hersheson, J., Mencacci, N.E., Davis, M., MacDonald, N., Trabzuni, D., Ryten, M., Pittman, A., Paudel, R., Kara, E., Fawcett, K., et al. (2013). Mutations in the autoregulatory domain of β-tubulin 4a cause hereditary dystonia. Ann. Neurol. *73*, 546–553.

Hirst, J., Barlow, L.D., Francisco, G.C., Sahlender, D. a, Seaman, M.N.J., Dacks, J.B., and Robinson, M.S. (2011). The fifth adaptor protein complex. PLoS Biol. *9*, e1001170.

Hirst, J., Irving, C., and Borner, G.H.H. (2013a). Adaptor Protein Complexes AP-4 and AP-5: New Players in Endosomal Trafficking and Progressive Spastic Paraplegia. Traffic *14*, 153–164.

Hirst, J., Borner, G.H.H., Edgar, J., Hein, M.Y., Mann, M., Buchholz, F., Antrobus, R., and Robinson, M.S. (2013b). Interaction between AP-5 and the hereditary spastic paraplegia

proteins SPG11 and SPG15. Mol. Biol. Cell *24*, 2558–2569.

Hodgkinson, C.A., Bohlega, S., Abu-Amero, S.N., Cupler, E., Kambouris, M., Meyer, B.F., and Bharucha, V.A. (2002). A novel form of autosomal recessive pure hereditary spastic paraplegia maps to chromosome 13q14. Neurology *59*, 1905–1909.

Holmes, G.L., and Shaywitz, B.A. (1977). Strumpell's pure familial spastic paraplegia: case study and review of the literature. J. Neurol. Neurosurg. Psychiatry *40*, 1003–1008.

Hortsch, M. (2000). Structural and Functional Evolution of the L1 Family: Are Four Adhesion Molecules Better Than One? Mol. Cell. Neurosci. *15*, 1–10.

Hu, Y.-J., and Imbalzano, A.N. (2016). Global gene expression profiling of JMJD6- and JMJD4-depleted mouse NIH3T3 fibroblasts. Sci. Data *3*, 160022.

Hu, C.-A.A., Bart Williams, D., Zhaorigetu, S., Khalil, S., Wan, G., and Valle, D. (2008a). Functional genomics and SNP analysis of human genes encoding proline metabolic enzymes. Amino Acids *35*, 655–664.

Hu, C.-A.A., Khalil, S., Zhaorigetu, S., Liu, Z., Tyler, M., Wan, G., and Valle, D. (2008b). Human Δ1-pyrroline-5-carboxylate synthase: function and regulation. Amino Acids *35*, 665–672.

Hu, J., Shibata, Y., Zhu, P.-P., Voss, C., Rismanchi, N., Prinz, W.A., Rapoport, T.A., Blackstone, C., Heiman-Patterson, T., Ming, L., et al. (2009). A class of dynamin-like GTPases involved in the generation of the tubular ER network. Cell *138*, 549–561.

Huang, H., Wang, H., and Figueiredo-Pereira, M.E. (2013). Regulating the ubiquitin/proteasome pathway via cAMP-signaling: neuroprotective potential. Cell Biochem. Biophys. *67*, 55–66.

Huber, M.D., Vesely, P.W., Datta, K., and Gerace, L. (2013). Erlins restrict SREBP activation in the ER and regulate cellular cholesterol homeostasis. J. Cell Biol. *203*, 427–436.

Ilgaz Aydinlar, E., Rolfs, A., Serteser, M., and Parman, Y. (2014). Mutation in FAM134B causing hereditary sensory neuropathy with spasticity in a Turkish family. Muscle and Nerve *49*, 774–775.

Inoue, K. (2005). PLP1-related inherited dysmyelinating disorders: Pelizaeus-Merzbacher disease and spastic paraplegia type 2. Neurogenetics *6*, 1–16.

Inoue, H., Baba, T., Sato, S., Ohtsuki, R., Takemori, A., Watanabe, T., Tagaya, M., and Tani, K. (2012). Roles of SAM and DDHD domains in mammalian intracellular phospholipase A1 KIAA0725p. Biochim. Biophys. Acta - Mol. Cell Res. *1823*, 930–939.

Inuzuka, M., Hayakawa, M., and Ingi, T. (2005). Serinc, an activity-regulated protein family, incorporates serine into membrane lipid synthesis. J. Biol. Chem. *280*, 35776–35783.

Iqbal, Z., Puttmann, L., Musante, L., Razzaq, A., Zahoor, M.Y., Hu, H., Wienker, T.F., Garshasbi, M., Fattahi, Z., Gilissen, C., et al. (2016). Missense variants in AIMP1 gene are implicated in autosomal recessive intellectual disability without neurodegeneration. Eur. J. Hum. Genet. *24*, 392–399.

Ishiura, H., Sako, W., Yoshida, M., Kawarai, T., Tanabe, O., Goto, J., Takahashi, Y., Date, H., Mitsui, J., Ahsan, B., et al. (2012). The TRK-Fused Gene Is Mutated in Hereditary Motor and Sensory Neuropathy with Proximal Dominant Involvement. Am. J. Hum. Genet. *91*, 320–329.

Ito, D., and Suzuki, N. (2007). Molecular pathogenesis of seipin/BSCL2-related motor neuron diseases. Ann. Neurol. *61*, 237–250.

Jacq, X., Kemp, M., Martin, N.M.B., and Jackson, S.P. (2013). Deubiquitylating Enzymes and DNA Damage Response Pathways. Cell Biochem Biophys *67*, 25–43.

Johri, A., and Beal, M.F. (2012). Mitochondrial dysfunction in neurodegenerative diseases. J. Pharmacol. Exp. Ther. *342*, 619–630.

Josifova, D.J., Monroe, G.R., Tessadori, F., Graaff, E. De, Mehta, S.G., Harakalova, M., Duran, K.J., Sanne, M.C., Jungbluth, H., Hoogenraad, C.C., et al. (2016). Heterozygous KIDINS220/ARMS nonsense variants cause spastic paraplegia, intellectual disability, nystagmus, and obesity. Hum. Mol. Genet.

Jouet, M., Rosenthal, A., Armstrong, G., MacFarlane, J., Stevenson, R., Paterson, J., Metzenberg, A., Ionasescu, V., Temple, K., and Kenwrick, S. (1994). X-linked spastic paraplegia (SPG1), MASA syndrome and X-linked hydrocephalus result from mutations in the L1 gene. Nat. Genet. *7*, 402–407.

Jovaisaite, V., Mouchiroud, L., and Auwerx, J. (2014). The mitochondrial unfolded protein response, a conserved stress response pathway with implications in health and disease. J. Exp. Biol. *217*, 137–143.

Kabashi, E., Brustein, E., Champagne, N., and Drapeau, P. (2011). Zebrafish models for the functional genomics of neurogenetic disorders. Biochim. Biophys. Acta - Mol. Basis Dis. *1812*, 335–345.

Kancheva, D., Chamova, T., Guergueltcheva, V., Mitev, V., Azmanov, D.N., Kalaydjieva, L., Tournev, I., and Jordanova, A. (2015). Mosaic dominant TUBB4A mutation in an inbred family with complicated hereditary spastic paraplegia. Mov. Disord. *30*, 854–858.

Kapiloff, M.S., Schillace, R. V, Westphal, a M., and Scott, J.D. (1999). mAKAP: an A-kinase anchoring protein targeted to the nuclear membrane of differentiated myocytes. J. Cell Sci. *112 (Pt 1*, 2725–2736.

Kara, E., Tucci, A., Manzoni, C., Lynch, D.S., Elpidorou, M., Bettencourt, C., Chelban, V., Manole, A., Hamed, S.A., Haridy, N.A., et al. (2016). Genetic and phenotypic characterization of complex hereditary spastic paraplegia. Brain *139*.

Kenwrick, S., Watkins, A., and De Angelis, E. (2000). Neural cell recognition molecule L1: relating biological complexity to human disease mutations. Hum. Mol. Genet. *9*, 879–886.

Khan, T.N., Klar, J., Tariq, M., Anjum Baig, S., Malik, N.A., Yousaf, R., Baig, S.M., and Dahl, N. (2014). Evidence for autosomal recessive inheritance in SPG3A caused by homozygosity for a novel ATL1 missense mutation. Eur. J. Hum. Genet. *22*, 1180–1184.

Klebe, S., Depienne, C., Gerber, S., Challe, G., Anheim, M., Charles, P., Fedirko, E., Lejeune, E., Cottineau, J., Brusco, A., et al. (2012). Spastic paraplegia gene 7 in patients with spasticity and/or optic neuropathy. Brain *135*, 2980–2993.

Klebe, S., Stevanin, G., and Depienne, C. (2015). Clinical and genetic heterogeneity in hereditary spastic paraplegias: From SPG1 to SPG72 and still counting. Rev. Neurol. (Paris). *171*, 505–530.

Kochergin, I.A., and Zakharova, M.N. (2016). The role of autophagy in neurodegenerative diseases. Neurochem. J. *10*, 7–18.

Kocoglu, C., Gundogdu, A., Kocaman, G., Kahraman-Koytak, P., Uluc, K., Kiziltan, G., Caglayan, A.O., Bilguvar, K., Vural, A., and Basak, A.N. (2018). Homozygous CAPN1 mutations causing a spastic-ataxia phenotype in 2 families. Neurol. Genet. *4*, e218.

Kritzer, M.D., Li, J., Passariello, C.L., Gayanilo, M., Thakur, H., Dayan, J., Dodge-Kafka, K., and Kapiloff, M.S. (2014). The scaffold protein muscle A-kinase anchoring protein β orchestrates cardiac myocyte hypertrophic signaling required for the development of heart failure. Circ. Heart Fail. *7*, 663–672.

Ku, C.S., Cooper, D.N., Polychronakos, C., Naidoo, N., Wu, M., and Soong, R. (2012). Exome sequencing: Dual role as a discovery and diagnostic tool. Ann. Neurol. *71*, 5–14.

Kumar, K.R., Blair, N.F., and Sue, C.M. (2015). An Update on the Hereditary Spastic Paraplegias: New Genes and New Disease Models. Mov. Disord. Clin. Pract. *2*, 213–223.

Kumar, K.R., Wali, G.M., Kamate, M., Wali, G., Minoche, A.E., Puttick, C., Pinese, M., Gayevskiy, V., Dinger, M.E., Roscioli, T., et al. (2016). Defining the genetic basis of early onset hereditary spastic paraplegia using whole genome sequencing. Neurogenetics *17*, 265–270.

Kuo, M.-L., Lee, M.B.-E., Tang, M., den Besten, W., Hu, S., Sweredoski, M.J., Hess, S., Chou, C.-M., Changou, C.A., Su, M., et al. (2016). PYCR1 and PYCR2 Interact and Collaborate with RRM2B to Protect Cells from Overt Oxidative Stress. Sci. Rep. *6*, 18846.

Landouré, G., Zhu, P., Lourenço, C.M., Johnson, J.O., Toro, C., Bricceno, K. V, Rinaldi, C., Meilleur, K.G., Sangaré, M., Pierson, T.M., et al. (2012). Hereditary spastic paraplegia type 43 (SPG43) is caused by mutation in C19orf12. Hum. Mutat. 1–16.

Lasiecka, Z.M., and Winckler, B. (2011). Mechanisms of polarized membrane trafficking in neurons - focusing in on endosomes. Mol. Cell. Neurosci. *48*, 278–287.

Lax, N.Z., Gnanapavan, S., Dowson, S.J., Alston, C.L., He, L., Polvikoski, T.M., Jaros, E., O'Donovan, D.G., Yarham, J.W., Turnbull, D.M., et al. (2013). Early-onset cataracts, spastic paraparesis, and ataxia caused by a novel mitochondrial tRNAGlu (MT-TE) gene mutation causing severe complex I deficiency: a clinical, molecular, and neuropathologic study. J. Neuropathol. Exp. Neurol. *72*, 164–175.

LeDoux, M. (2005). CHAPTER A2 – Animal Models and the Science of Movement Disorders. In Animal Models of Movement Disorders, pp. 13–31.

Lee, C.Y., Choi, J.-W., Shin, S., Lee, J., Seo, H.-H., Lim, S., Lee, S., Joo, H.-C., Kim, S.W., and Hwang, K.-C. (2017a). Interaction of small G protein signaling modulator 3 with connexin 43 contributes to myocardial infarction in rat hearts. Biochem. Biophys. Res. Commun. *491*, 429–435.

Lee, I.K., Kim, K.S., Kim, H., Lee, J.Y., Ryu, C.H., Chun, H.J., Lee, K.U., Lim, Y., Kim, Y.H., Huh, P.W., et al. (2004). MAP, a protein interacting with a tumor suppressor, merlin, through the run domain. Biochem. Biophys. Res. Commun. *325*, 774–783.

Lee, J.R., Srour, M., Kim, D., Hamdan, F.F., Lim, S.H., Brunel-Guitton, C., Décarie, J.C., Rossignol, E., Mitchell, G.A., Schreiber, A., et al. (2015a). De novo mutations in the motor domain of KIF1A cause cognitive impairment, spastic paraparesis, axonal neuropathy, and cerebellar atrophy. Hum. Mutat. *36*, 69–78.

Lee, J.Y.W., Hsu, C.-K., Michael, M., Nanda, A., Liu, L., McMillan, J.R., Pourreyron, C., Takeichi, T., Tolar, J., Reid, E., et al. (2017b). Large Intragenic Deletion in DSTYK Underlies Autosomal-Recessive Complicated Spastic Paraparesis, SPG23. Am. J. Hum. Genet. *100*, 364–370.

Lee, S.-W., Won, J.-Y., Yang, J., Lee, J., Kim, S.-Y., Lee, E.J., and Kim, H.-S. (2015b). AKAP6 inhibition impairs myoblast differentiation and muscle regeneration: Positive loop between AKAP6 and myogenin. Sci. Rep. *5*, 16523.

Lerer, I., Sagi, M., Meiner, V., Cohen, T., Zlotogora, J., and Abeliovich, D. (2005). Deletion of the ANKRD15 gene at 9p24.3 causes parent-of-origin-dependent inheritance of familial cerebral palsy. Hum. Mol. Genet. *14*, 3911–3920.

Lim, L., Shui, G., and Wenk, M.R. (2012). Lipid Metabolism in Neurodegenerative Diseases. In Lipidomics: Technologies and Applications, pp. 269–296.

Lin, P., Li, J., Liu, Q., Mao, F., Li, J., Qiu, R., Hu, H., Song, Y., Yang, Y., Gao, G., et al. (2008). A Missense Mutation in SLC33A1, which Encodes the Acetyl-CoA Transporter, Causes Autosomal-Dominant Spastic Paraplegia (SPG42). Am. J. Hum. Genet. *83*, 752–759.

Lindholm, D., Wootz, H., and Korhonen, L. (2006). ER stress and neurodegenerative diseases. Cell Death Differ. *13*, 385–392.

Lindstrand, A., Frangakis, S., Carvalho, C.M.B., Richardson, E.B., McFadden, K.A., Willer, J.R., Pehlivan, D., Liu, P., Pediaditakis, I.L., Sabo, A., et al. (2016). Copy-Number Variation Contributes to the Mutational Load of Bardet-Biedl Syndrome. Am. J. Hum. Genet. *99*, 318–336.

Liu, A., and Niswander, L.A. (2005). Bone morphogenetic protein signalling and vertebrate nervous system development. Nat. Rev. Neurosci. *6*, 945–954.

Liu, Y., and Zhu, X. (2017). Endoplasmic reticulum-mitochondria tethering in neurodegenerative diseases. Transl. Neurodegener. *6*, 21.

Liu, J., Liu, M.C., and Wang, K.K.W. (2008). Calpain in the CNS: From Synaptic Function to

Neurotoxicity. Sci. Signal. *1*, re1-re1.

Liu, Y., Fallon, L., Lashuel, H.A., Liu, Z., and Lansbury, P.T. (2002). The UCH-L1 gene encodes two opposing enzymatic activities that affect alpha-synuclein degradation and Parkinson's disease susceptibility. Cell *111*, 209–218.

Liu, Y., Lee, J.W., and Ackerman, S.L. (2015). Mutations in the microtubule-associated protein 1A (Map1a) gene cause Purkinje cell degeneration. J. Neurosci. *35*, 4587–4598.

Lohmann, K., Wilcox, R.A., Winkler, S., Ramirez, A., Rakovic, A., Park, J.-S., Arns, B., Lohnau, T., Groen, J., Kasten, M., et al. (2013). Whispering dysphonia (DYT4 dystonia) is caused by a mutation in the TUBB4 gene. Ann. Neurol. *73*, 537–545.

Lorrain, M. (1898). Contribution à l'étude de la paraplégie spasmodique familiale: travail de la clinique des maladies du système nerveux à la Salpêtrière.

Lossos, A., Stümpfig, C., Stevanin, G., Gaussen, M., Zimmerman, B.-E., Mundwiller, E., Asulin, M., Chamma, L., Sheffer, R., Misk, A., et al. (2015). Fe/S protein assembly gene IBA57 mutation causes hereditary spastic paraplegia. Neurology *84*, 659–667.

Loureiro, J.L., Miller-Fleming, L., Thieleke-Matos, C., Magalhães, P., Cruz, V.T., Coutinho, P., Sequeiros, J., and Silveira, I. (2009). Novel SPG3A and SPG4 mutations in dominant spastic paraplegia families. Acta Neurol. Scand. *119*, 113–118.

Lu, J., Rashid, F., and Byrne, P.C. (2006). The hereditary spastic paraplegia protein spartin localises to mitochondria. J. Neurochem. *98*, 1908–1919.

Luo, Y., Du, J., Zhan, Z., Chen, C., Wang, J., Hu, Y., Hu, Z., Xia, K., Tang, B., and Shen, L. (2013). A diagnostic gene chip for hereditary spastic paraplegias. Brain Res. Bull. *97*, 112–118.

Lynch, D.S., Koutsis, G., Tucci, A., Panas, M., Baklou, M., Breza, M., Karadima, G., and Houlden, H. (2016). Hereditary spastic paraplegia in Greece: characterisation of a previously unexplored population using next-generation sequencing. Eur. J. Hum. Genet. *24*, 857–863.

Lynes, E.M., and Simmen, T. (2011). Urban planning of the endoplasmic reticulum (ER): How diverse mechanisms segregate the many functions of the ER. Biochim. Biophys. Acta - Mol. Cell Res. *1813*, 1893–1905.

Lynex, C.N., Carr, I.M., Leek, J.P., Achuthan, R., Mitchell, S., Maher, E.R., Woods, C.G., Bonthon, D.T., and Markham, A.F. (2004). Homozygosity for a missense mutation in the 67 kDa isoform of glutamate decarboxylase in a family with autosomal recessive spastic cerebral palsy: parallels with Stiff-Person Syndrome and other movement disorders. BMC Neurol. *4*, 20.

Macedo-Souza, L.I., Kok, F., Santos, S., Licinio, L., Lezirovitz, K., Nascimento, R.M.P., Bueno, C., Martyn, M., Leão, E.K.E.A., and Zatz, M. (2008). Reevaluation of a large family defines a new locus for X-linked recessive pure spastic paraplegia (SPG34) on chromosome Xq25. Neurogenetics *9*, 225–226.

Maday, S., Twelvetrees, A.E., Moughamian, A.J., and Holzbaur, E.L.F. (2014). Axonal Transport: Cargo-Specific Mechanisms of Motility and Regulation. Neuron *84*, 292–309.

Magen, D., Georgopoulos, C., Bross, P., Ang, D., Segev, Y., Goldsher, D., Nemirovski, A., Shahar, E., Ravid, S., Luder, A., et al. (2008). Mitochondrial Hsp60 Chaperonopathy Causes an Autosomal-Recessive Neurodegenerative Disorder Linked to Brain Hypomyelination and Leukodystrophy. Am. J. Hum. Genet. *83*, 30–42.

Maier, A., Klopocki, E., Horn, D., Tzschach, A., Holm, T., Meyer, R., and Meyer, T. (2014). *De novo* partial deletion in *GRID2* presenting with complicated spastic paraplegia. Muscle Nerve *49*, 289–292.

Mannan, A.U., Krawen, P., Sauter, S.M., Boehm, J., Chronowska, A., Paulus, W., Neesen, J., and Engel, W. (2006). ZFYVE27 (SPG33), a novel spastin-binding protein, is mutated in hereditary spastic paraplegia. Am. J. Hum. Genet. *79*, 351–357.

Manole, A., Männikkö, R., Hanna, M.G., Kullmann, D.M., and Houlden, H. (2017). De novo *KCNA2* mutations cause hereditary spastic paraplegia. Ann. Neurol. *81*, 326–328.

Martin, E., Yanicostas, C., Rastetter, A., Alavi Naini, S.M., Maouedj, A., Kabashi, E., Rivaud-Péchoux, S., Brice, A., Stevanin, G., and Soussi-Yanicostas, N. (2012). Spatacsin and spastizin act in the same pathway required for proper spinal motor neuron axon outgrowth in zebrafish. Neurobiol. Dis. *48*, 299–308.

Martin, E., Schüle, R., Smets, K., Rastetter, A., Boukhris, A., Loureiro, J.L., Gonzalez, M.A., Mundwiller, E., Deconinck, T., Wessner, M., et al. (2013). Loss of function of glucocerebrosidase GBA2 is responsible for motor neuron defects in hereditary spastic paraplegia. Am. J. Hum. Genet. *92*, 238–244.

Matsuto, M., Kano, F., and Murata, M. (2015). Reconstitution of the targeting of Rab6A to the Golgi apparatus in semi-intact HeLa cells: A role of BICD2 in stabilizing Rab6A on Golgi membranes and a concerted role of Rab6A/BICD2 interactions in Golgi-to-ER retrograde transport. Biochim. Biophys. Acta - Mol. Cell Res. *1853*, 2592–2609.

McCartney, S., Little, B.M., Langeberg, L.K., and Scott, J.D. (1995). Cloning and characterization of A-kinase anchor protein 100 (AKAP100). A protein that targets A-kinase to the sarcoplasmic reticulum. J. Biol. Chem. *270*, 9327–9333.

McCaughey, J., Miller, V.J., Stevenson, N.L., Brown, A.K., Budnik, A., Heesom, K.J., Alibhai, D., and Stephens, D.J. (2016). TFG Promotes Organization of Transitional ER and Efficient Collagen Secretion. Cell Rep. *15*, 1648–1659.

Meijer, I.A., Cossette, P., Roussel, J., Benard, M., Toupin, S., and Rouleau, G.A. (2004). A novel locus for pure recessive hereditary spastic paraplegia maps to 10q22.1-10q24.1. Ann. Neurol. *56*, 579–582.

Melo, U.S., Macedo-Souza, L.I., Figueiredo, T., Muotri, A.R., Gleeson, J.G., Coux, G., Armas, P., Calcaterra, N.B., Kitajima, J.P., Amorim, S., et al. (2015). Overexpression of KLC2 due to

a homozygous deletion in the non-coding region causes SPOAN syndrome. Hum. Mol. Genet. *24*, 6877–6885.

Michel, J.J.C., Townley, I.K., Dodge-Kafka, K.L., Zhang, F., Kapiloff, M.S., and Scott, J.D. (2005). Spatial restriction of PDK1 activation cascades by anchoring to mAKAPα. Mol. Cell *20*, 661–672.

Micheli, V., Camici, M., Tozzi, M.G., Ipata, P.L., Sestini, S., Bertelli, M., and Pompucci, G. (2011). Neurological disorders of purine and pyrimidine metabolism. Curr. Top. Med. Chem. *11*, 923–947.

Miyatake, S., Osaka, H., Shiina, M., Sasaki, M., Takanashi, J. -i., Haginoya, K., Wada, T., Morimoto, M., Ando, N., Ikuta, Y., et al. (2014). Expanding the phenotypic spectrum of TUBB4A-associated hypomyelinating leukoencephalopathies. Neurology *82*, 2230–2237.

Momeni, P., Schymick, J., Jain, S., Cookson, M.R., Cairns, N.J., Greggio, E., Greenway, M.J., Berger, S., Pickering-Brown, S., Chiò, A., et al. (2006). Analysis of IFT74 as a candidate gene for chromosome 9p-linked ALS-FTD. BMC Neurol. *6*, 44.

Montenegro, G., Rebelo, A.P., Connell, J., Allison, R., Babalini, C., D'Aloia, M., Montieri, P., Schüle, R., Ishiura, H., Price, J., et al. (2012). Mutations in the ER-shaping protein reticulon 2 cause the axon-degenerative disorder hereditary spastic paraplegia type 12. J. Clin. Invest. *122*, 538–544.

Morais, S., Raymond, L., Mairey, M., Coutinho, P., Brandão, E., Ribeiro, P., Loureiro, J.L., Sequeiros, J., Brice, A., Alonso, I., et al. (2017). Massive sequencing of 70 genes reveals a myriad of missing genes or mechanisms to be uncovered in hereditary spastic paraplegias. Eur. J. Hum. Genet. *25*, 1217–1228.

Murakami, Y., Tawamie, H., Maeda, Y., Bü Ttner, C., Buchert, R., Radwan, F., Schaffer, S., Sticht, H., Aigner, M., Reis, A., et al. (2014). Null Mutation in PGAP1 Impairing Gpi-Anchor Maturation in Patients with Intellectual Disability and Encephalopathy. PLoS Genet. *10*, 1–8.

Murmu, R.P., Martin, E., Rastetter, A., Esteves, T., Muriel, M.P., El Hachimi, K.H., Denora, P.S., Dauphin, A., Fernandez, J.C., Duyckaerts, C., et al. (2011). Cellular distribution and subcellular localization of spatacsin and spastizin, two proteins involved in hereditary spastic paraplegia. Mol. Cell. Neurosci. *47*, 191–202.

Musacchio, T., Zaum, A.-K., Üçeyler, N., Sommer, C., Pfeifroth, N., Reiners, K., Kunstmann, E., Volkmann, J., Rost, S., and Klebe, S. (2017). ALS and MMN mimics in patients with BSCL2 mutations: the expanding clinical spectrum of SPG17 hereditary spastic paraplegia. J. Neurol. *264*, 11–20.

Nakayama, T., Al-Maawali, A., El-Quessny, M., Rajab, A., Khalil, S., Stoler, J.M., Tan, W.-H., Nasir, R., Schmitz-Abe, K., Hill, R.S., et al. (2015). Mutations in PYCR2, Encoding Pyrroline-5-Carboxylate Reductase 2, Cause Microcephaly and Hypomyelination. Am. J. Hum. Genet. *96*, 709–719.

Namekawa, M., Muriel, M.-P., Janer, A., Latouche, M., Dauphin, A., Debeir, T., Martin, E., Duyckaerts, C., Prigent, A., Depienne, C., et al. (2007). Mutations in the SPG3A gene encoding the GTPase atlastin interfere with vesicle trafficking in the ER/Golgi interface and Golgi morphogenesis. Mol. Cell. Neurosci. *35*, 1–13.

Nardi-Schreiber, A., Sapir, G., Gamliel, A., Kakhlon, O., Sosna, J., Gomori, J.M., Meiner, V., Lossos, A., and Katz-Brull, R. (2017). Defective ATP breakdown activity related to an ENTPD1 gene mutation demonstrated using (31)P NMR spectroscopy. Chem. Commun. (Camb). *53*, 9121–9124.

Neveling, K., Martinez-Carrera, L.A., Hölker, I., Heister, A., Verrips, A., Hosseini-Barkooie, S.M., Gilissen, C., Vermeer, S., Pennings, M., Meijer, R., et al. (2013). Mutations in BICD2, which encodes a golgin and important motor adaptor, cause congenital autosomal-dominant spinal muscular atrophy. Am. J. Hum. Genet. *92*, 946–954.

Ng, P.C., and Henikoff, S. (2006). Predicting the Effects of Amino Acid Substitutions on Protein Function. Annu. Rev. Genomics Hum. Genet. *7*, 61–80.

Nolden, M., Ehses, S., Koppen, M., Bernacchia, A., Rugarli, E.I., and Langer, T. (2005). The m-AAA Protease Defective in Hereditary Spastic Paraplegia Controls Ribosome Assembly in Mitochondria. Cell *123*, 277–289.

Noreau, A., Dion, P.A., and Rouleau, G.A. (2014). Molecular aspects of hereditary spastic paraplegia. Exp. Cell Res. *325*.

Novarino, G., Fenstermaker, a G., Zaki, M.S., Hofree, M., Silhavy, J.L., Heiberg, a D., Abdellateef, M., Rosti, B., Scott, E., Mansour, L., et al. (2014). Exome sequencing links corticospinal motor neuron disease to common neurodegenerative disorders. Science (80-.). *343*, 506–511.

Oates, E.C., Rossor, A.M., Hafezparast, M., Gonzalez, M., Speziani, F., Macarthur, D.G., Lek, M., Cottenie, E., Scoto, M., Foley, A.R., et al. (2013). Mutations in BICD2 cause dominant congenital spinal muscular atrophy and hereditary spastic paraplegia. Am. J. Hum. Genet. *92*, 965–973.

Ohmi, Y., Tajima, O., Ohkawa, Y., Yamauchi, Y., Sugiura, Y., Furukawa, K., and Furukawa, K. (2011). Gangliosides are essential in the protection of inflammation and neurodegeneration via maintenance of lipid rafts: elucidation by a series of ganglioside-deficient mutant mice. J. Neurochem. *116*, 926–935.

Orlacchio, A., Kawarai, T., Gaudiello, F., St. George-Hyslop, P.H., Floris, R., and Bernardi, G. (2005). New locus for hereditary spastic paraplegia maps to chromosome 1p31.1-1p21.1. Ann. Neurol. *58*, 423–429.

Orlacchio, A., Patrono, C., Gaudiello, F., Rocchi, C., Moschella, V., Floris, R., Bernardi, G., and Kawarai, T. (2008). Silver syndrome variant of hereditary spastic paraplegia: A locus to 4p and allelism with SPG4. Neurology *70*, 1959–1966.

Orso, G., Pendin, D., Liu, S., Tosetto, J., Moss, T.J., Faust, J.E., Micaroni, M., Egorova, A., Martinuzzi, A., McNew, J.A., et al. (2009). Homotypic fusion of ER membranes requires the dynamin-like GTPase Atlastin. Nature *460*, 978–983.

Orthmann-Murphy, J.L., Salsano, E., Abrams, C.K., Bizzi, A., Uziel, G., Freidin, M.M., Lamantea, E., Zeviani, M., Scherer, S.S., and Pareyson, D. (2009). Hereditary spastic paraplegia is a novel phenotype for GJA12/GJC2 mutations. Brain *132*, 426–438.

Osses, N., and Henríquez, J.P. (2014). Bone morphogenetic protein signaling in vertebrate motor neurons and neuromuscular communication. Front. Cell. Neurosci. *8*, 453.

Oudega, M., Touri, F., Deenen, M.G., Riederer, B.M., and Marani, E. (1998). Microtubule-associated protein 1a is involved in the early development of the rat spinal cord. Neurosci. Lett. *246*, 81–84.

Oz-Levi, D., Ben-Zeev, B., Ruzzo, E.K., Hitomi, Y., Gelman, A., Pelak, K., Anikster, Y., Reznik-Wolf, H., Bar-Joseph, I., Olender, T., et al. (2012). Mutation in TECPR2 reveals a role for autophagy in hereditary spastic paraparesis. Am. J. Hum. Genet. *91*, 1065–1072.

Oz-Levi, D., Gelman, A., Elazar, Z., and Lancet, D. (2013). TECPR2:A new autophagy link for neurodegeneration. Autophagy *9*, 801–802.

Ozes, B., Karagoz, N., Schüle, R., Rebelo, A., Sobrido, M.J., Harmuth, F., Synofzik, M., Pascual, S.I.P., Colak, M., Ciftci-Kavaklioglu, B., et al. (2017). PLA2G6 mutations associated with a continuous clinical spectrum from neuroaxonal dystrophy to hereditary spastic paraplegia. Clin. Genet. 1–6.

Paralkar, V.M.V., Grgurevic, L., Simic, P., Ryskin, M., Hong, Q., Brault, A.L. AL, Lubeski, C., Brenner, M.M.B., Pauk□, M., Riccardi, K.K.A., et al. (2010). A novel role of bone morphogenetic protein-6 (BMP-6) as an endocrine regulator of bone and glucose homeostasis. Bone *47*, S57.

Pare, G.C., Easlick, J.L., Mislow, J.M., McNally, E.M., and Kapiloff, M.S. (2005). Nesprin-1α contributes to the targeting of mAKAP to the cardiac myocyte nuclear envelope. Exp. Cell Res. *303*, 388–399.

Park, J., Lee, B.-S., Choi, J.-K., Means, R.E., Choe, J., and Jung, J.U. (2002). Herpesviral protein targets a cellular WD repeat endosomal protein to downregulate T lymphocyte receptor expression. Immunity *17*, 221–233.

Park, S.G., Kang, Y.S., Kim, J.Y., Lee, C.S., Ko, Y.G., Lee, W.J., Lee, K.-U., Yeom, Y. Il, and Kim, S. (2006). Hormonal activity of AIMP1/p43 for glucose homeostasis. Proc. Natl. Acad. Sci. U. S. A. *103*, 14913–14918.

Park, S.G., Schimmel, P., and Kim, S. (2008). Aminoacyl tRNA synthetases and their connections to disease. Proc. Natl. Acad. Sci. U. S. A. *105*, 11043–11049.

Park, S.H., Zhu, P.-P., L., P.R., Blackstone, C., Parker, R.L., and Blackstone, C. (2010). Hereditary spastic paraplegia proteins REEP1, spastin, and atlastin-1 coordinate microtubule

interactions with the tubular ER network. J. Clin. Invest. *120*, 1097–1110.

Parodi, L., Fenu, S., Stevanin, G., and Durr, A. (2017). Hereditary spastic paraplegia: More than an upper motor neuron disease. Rev. Neurol. (Paris). 1–9.

Passariello, C.L., Li, J., Dodge-Kafka, K., and Kapiloff, M.S. (2015). mAKAP – A Master Scaffold for Cardiac Remodeling. J. Cardiovasc. Pharmacol. *65*, 218–225.

Patel, H., Cross, H., Proukakis, C., Hershberger, R., Bork, P., Ciccarelli, F.D., Patton, M.A., McKusick, V.A., and Crosby, A.H. (2002). SPG20 is mutated in Troyer syndrome, an hereditary spastic paraplegia. Nat. Genet. *31*, 347.

Patten, S.A., Armstrong, G.A.B., Lissouba, A., Kabashi, E., Parker, J.A., and Drapeau, P. (2014). Fishing for causes and cures of motor neuron disorders. Dis. Model. Mech. *7*, 799–809.

Pehar, M., Jonas, M.C., Hare, T.M., and Puglielli, L. (2012). SLC33A1/AT-1 protein regulates the induction of autophagy downstream of IRE1/XBP1 pathway. J. Biol. Chem. *287*, 29921–29930.

La Piana, R., Vanasse, M., Brais, B., and Bernard, G. (2015). Myelination Delay and Allan-Herndon-Dudley Syndrome Caused by a Novel Mutation in the SLC16A2 Gene. J. Child Neurol. *30*, 1371–1374.

Pierson, T.M., Adams, D., Bonn, F., Martinelli, P., Cherukuri, P.F., Teer, J.K., Hansen, N.F., Cruz, P., Mullikin, J.C., Blakesley, R.W., et al. (2011). Whole-exome sequencing identifies homozygous AFG3L2 mutations in a spastic ataxia-neuropathy syndrome linked to mitochondrial m-AAA proteases. PLoS Genet. *7*.

Poswar, F. de O., da Fonseca, R.M.T., de Albuquerque, L.C.P., Zhou, Q., Jardim, L.B., Monte, T.L., Aksentijevich, I., and Saute, J.A.M. (2016). Adenosine deaminase 2 deficiency presenting as spastic paraplegia and systemic vasculitis. J. Neurol. *263*, 818–820.

Quarles, R.H. (2007). Myelin-associated glycoprotein (MAG): past, present and beyond. J. Neurochem. *0*, 070214184024009–???

Quevillon, S., Agou, F., Robinson, J.C., and Mirande, M. (1997). The p43 component of the mammalian multi-synthetase complex is likely to be the precursor of the endothelial monocyte-activating polypeptide II cytokine. J. Biol. Chem. *272*, 32573–32579.

Rababa'H, A., Craft, J.W., Wijaya, C.S., Atrooz, F., Fan, Q., Singh, S., Guillory, A.N., Katsonis, P., Lichtarge, O., and McConnell, B.K. (2013). Protein kinase A and phosphodiesterase-4D3 binding to coding polymorphisms of cardiac muscle anchoring protein (mAKAP). J. Mol. Biol. *425*, 3277–3288.

Rahman, M.S., Akhtar, N., Jamil, H.M., Banik, R.S., and Asaduzzaman, S.M. (2015). TGF-β/BMP signaling and other molecular events: regulation of osteoblastogenesis and bone formation. Bone Res. *3*, 15005.

Raiborg, C., Wenzel, E.M., and Stenmark, H. (2015). ER-endosome contact sites: molecular

compositions and functions. EMBO J. *34*, 1848–1858.

Rainier, S., and Fink, J.K. (2005). CHAPTER L2 – Hereditary Spastic Paraplegia: Clinical Features and Animal Models. In Animal Models of Movement Disorders, pp. 687–690.

Rainier, S., Chai, J.-H., Tokarz, D., Nicholls, R.D., and Fink, J.K. (2003). NIPA1 Gene Mutations Cause Autosomal Dominant Hereditary Spastic Paraplegia (SPG6). Am. J. Hum. Genet. *73*, 967–971.

Rainier, S., Bui, M., Mark, E., Thomas, D., Tokarz, D., Ming, L., Delaney, C., Richardson, R.J., Albers, J.W., Matsunami, N., et al. (2008). Neuropathy Target Esterase Gene Mutations Cause Motor Neuron Disease. Am. J. Hum. Genet. *82*, 780–785.

Ravera, S., and Panfoli, I. (2015). Role of myelin sheath energy metabolism in neurodegenerative diseases. Neural Regen. Res. *10*, 1570–1571.

Reid, E., Kloos, M., Ashley-Koch, A., Hughes, L., Bevan, S., Svenson, I.K., Graham, F.L., Gaskell, P.C., Dearlove, A., Pericak-Vance, M.A., et al. (2002). A kinesin heavy chain (KIF5A) mutation in hereditary spastic paraplegia (SPG10). Am. J. Hum. Genet. *71*, 1189–1194.

Renvoisé, B., Chang, J., Singh, R., Yonekawa, S., FitzGibbon, E.J., Mankodi, A., Vanderver, A., Schindler, A., Toro, C., Gahl, W.A., et al. (2014). Lysosomal abnormalities in hereditary spastic paraplegia types SPG15 and SPG11. Ann. Clin. Transl. Neurol. *1*, 379–389.

Reversade, B., Escande-Beillard, N., Dimopoulou, A., Fischer, B., Chng, S.C., Li, Y., Shboul, M., Tham, P.-Y.Y., Kayserili, H., Al-Gazali, L., et al. (2009). Mutations in PYCR1 cause cutis laxa with progeroid features. Nat. Genet. *41*, 1016–1021.

Rinaldi, C., Schmidt, T., Situ, A.J., Johnson, J.O., Lee, P.R., Chen, K., Bott, L.C., Fadó, R., Harmison, G.H., Parodi, S., et al. (2015). Mutation in CPT1C Associated With Pure Autosomal Dominant Spastic Paraplegia. JAMA Neurol. *72*, 561–570.

Røsok, Ø., Pedeutour, F., Odeberg, J., Lundeberg, J., and Aasheim, H.-C. (2000). The C1orf9 Gene Encodes a Putative Transmembrane Member of a Novel Protein Family. Biochem. Biophys. Res. Commun. *267*, 855–862.

Ross, C.A., and Truant, R. (2017). A unifying mechanism in neurodegeneration. Nature *541*, 34–35.

Ruano, L., Melo, C., Silva, M.C., and Coutinho, P. (2014). The global epidemiology of hereditary ataxia and spastic paraplegia: A systematic review of prevalence studies. Neuroepidemiology *42*, 174–183.

Rugarli, E.I., and Langer, T. (2006). Translating m-AAA protease function in mitochondria to hereditary spastic paraplegia. Trends Mol. Med. *12*, 262–269.

Rydning, S.L., Backe, P.H., Sousa, M.M.L.L., Iqbal, Z., Øye, A.-M.M., Sheng, Y., Yang, M., Lin, X., Slupphaug, G., Nordenmark, T.H., et al. (2017). Novel UCHL1 mutations reveal new insights into ubiquitin processing. Hum. Mol. Genet. *26*, 1031–1040.

Rydning, S.L., Dudesek, A., Rimmele, F., Funke, C., Krüger, S., Biskup, S., Vigeland, M.D.,

Hjorthaug, H.S., Sejersted, Y., Tallaksen, C., et al. (2018). A novel heterozygous variant in *ERLIN2* causes autosomal dominant pure hereditary spastic paraplegia. Eur. J. Neurol.

Salinas, S., Carazo-Salas, R.E., Proukakis, C., Schiavo, G., and Warner, T.T. (2007). Spastin and Microtubules: Functions in Health and Disease. J. Neurosci. Res. *85*, 2778–2782.

Salinas, S., Proukakis, C., Crosby, A., and Warner, T.T. (2008). Hereditary spastic paraplegia: clinical features and pathogenetic mechanisms. Lancet Neurol. *7*, 1127–1138.

Salo, V.T., Belevich, I., Li, S., Karhinen, L., Vihinen, H., Vigouroux, C., Magré, J., Thiele, C., Hölttä-Vuori, M., Jokitalo, E., et al. (2016). Seipin regulates ER–lipid droplet contacts and cargo delivery. EMBO J. *35*, 2699–2716.

Sambuughin, N., Goldfarb, L.G., Sivtseva, T.M., Davydova, T.K., Vladimirtsev, V.A., Osakovskiy, V.L., Danilova, A.P., Nikitina, R.S., Ylakhova, A.N., Diachkovskaya, M.P., et al. (2015). Adult-onset autosomal dominant spastic paraplegia linked to a GTPase-effector domain mutation of dynamin 2. BMC Neurol. *15*, 223.

Sánchez-Ferrero, E., Coto, E., Beetz, C., Gámez, J., Corao, A., Díaz, M., Esteban, J., del Castillo, E., Moris, G., Infante, J., et al. (2013). SPG7 mutational screening in spastic paraplegia patients supports a dominant effect for some mutations and a pathogenic role for p.A510V. Clin. Genet. *83*, 257–262.

Saugier-Veber, P., Munnich, A., Bonneau, D., Rozet, J.-M., Merrer, M., Gil, R., and Boespflug-Tanguy, O. (1994). X-linked spastic paraplegia and Pelizaeus-Merzbacher disease are allelic disorders at the proteolipid protein locus. Nat. Genet. *6*, 257–262.

Schöls, L., Rattay, T.W., Martus, P., Meisner, C., Baets, J., Fischer, I., Jägle, C., Fraidakis, M.J., Martinuzzi, A., Saute, J.A., et al. (2017). Hereditary spastic paraplegia type 5: natural history, biomarkers and a randomized controlled trial. Brain *140*, 3112–3127.

Schule, R., Bonin, M., Durr, A., Forlani, S., Sperfeld, A.D., Klimpe, S., Mueller, J.C., Seibel, A., van de Warrenburg, B.P., Bauer, P., et al. (2009). Autosomal dominant spastic paraplegia with peripheral neuropathy maps to chr12q23-24. Neurology *72*, 1893–1898.

Schüle, R., Wiethoff, S., Martus, P., Karle, K.N., Otto, S., Klebe, S., Klimpe, S., Gallenmüller, C., Kurzwelly, D., Henkel, D., et al. (2016). Hereditary spastic paraplegia: Clinicogenetic lessons from 608 patients. Ann. Neurol. *79*, 646–658.

Schuurs-Hoeijmakers, J.H.M., Geraghty, M.T., Kamsteeg, E.J., Ben-Salem, S., De Bot, S.T., Nijhof, B., Van De Vondervoort, I.I.G.M., Van Der Graaf, M., Nobau, A.C., Otte-Höller, I., et al. (2012). Mutations in DDHD2, encoding an intracellular phospholipase A1, cause a recessive form of complex hereditary spastic paraplegia. Am. J. Hum. Genet. *91*, 1073–1081.

Schwartz, C.E., May, M.M., Carpenter, N.J., Rogers, R.C., Martin, J., Bialer, M.G., Ward, J., Sanabria, J., Marsa, S., Lewis, J.A., et al. (2005). Allan-Herndon-Dudley Syndrome and the Monocarboxylate Transporter 8 (MCT8) Gene. Am. J. Hum. Genet. *77*, 41–53.

Schwarz, D.S., and Blower, M.D. (2016). The endoplasmic reticulum: Structure, function and

response to cellular signaling. Cell. Mol. Life Sci. *73*, 79–94.

Schwarz, G.A., and Liu, C.-N. (1956). Hereditary (familial) spastic paraplegia; further clinical and pathologic observations. AMA Arch. Neurol. Psychiatry *75*, 144–162.

Schwarz, L.A., and Patrick, G.N. (2012). Ubiquitin-dependent endocytosis, trafficking and turnover of neuronal membrane proteins. Mol. Cell. Neurosci. *49*, 387–393.

Seeligmuller A. (1876). Sklerose der Seitenstränge des Rückenmarks bei vier Kindern derselben Familie. Dtsch Med Wschr. *2*, 185–186.

Sha, Z., Sha, L., Li, W., Dou, W., Shen, Y., Wu, L., and Xu, Q. (2015). Exome sequencing identifies SUCO mutations in mesial temporal lobe epilepsy. Neurosci. Lett. *591*, 149–154.

Sheftel, A.D., Wilbrecht, C., Stehling, O., Niggemeyer, B., Elsasser, H.-P., Muhlenhoff, U., and Lill, R. (2012). The human mitochondrial ISCA1, ISCA2, and IBA57 proteins are required for [4Fe-4S] protein maturation. Mol. Biol. Cell *23*, 1157–1166.

Shibata, Y., Shemesh, T., Prinz, W.A., Palazzo, A.F., Kozlov, M.M., Rapoport, T.A., Hattori, S., Ogawa-Goto, K., and Hawes, C. (2010). Mechanisms determining the morphology of the peripheral ER. Cell *143*, 774–788.

Shimazaki, H., Takiyama, Y., Ishiura, H., Sakai, C., Matsushima, Y., Hatakeyama, H., Honda, J., Sakoe, K., Naoi, T., Namekawa, M., et al. (2012). A homozygous mutation of C12orf65 causes spastic paraplegia with optic atrophy and neuropathy (SPG55). J. Med. Genet. *49*, 777–784.

Shimazaki, H., Honda, J., Naoi, T., Namekawa, M., Nakano, I., Yazaki, M., Nakamura, K., Yoshida, K., Ikeda, S., Ishiura, H., et al. (2014). Autosomal-recessive complicated spastic paraplegia with a novel lysosomal trafficking regulator gene mutation. J. Neurol. Neurosurg. Psychiatry *85*, 1024–1028.

Shoukier, M., Neesen, J., Sauter, S.M., Argyriou, L., Doerwald, N., Pantakani, D.K., and Mannan, A.U. (2009). Expansion of mutation spectrum, determination of mutation cluster regions and predictive structural classification of SPAST mutations in hereditary spastic paraplegia. Eur. J. Hum. Genet. *17*, 187–194.

Silva, M.C., Coutinho, P., Pinheiro, C.D., Neves, J.M., and Serrano, P. (1997). Hereditary ataxias and spastic paraplegias: Methodological aspects of a prevalence study in Portugal. J. Clin. Epidemiol. *50*, 1377–1384.

Simpson, M.A., Cross, H., Proukakis, C., Pryde, A., Hershberger, R., Chatonnet, A., Patton, M.A., and Crosby, A.H. (2003). Maspardin Is Mutated in Mast Syndrome, a Complicated Form of Hereditary Spastic Paraplegia Associated with Dementia. Am. J. Hum. Genet. *73*, 1147–1156.

Singh, G., and Davenport, A.P. (2006). Neuropeptide B and W: neurotransmitters in an emerging G-protein-coupled receptor system. Br. J. Pharmacol. *148*, 1033–1041.

Slabicki, M., Theis, M., Krastev, D.B., Samsonov, S., Mundwiller, E., Junqueira, M.,

Paszkowski-Rogacz, M., Teyra, J., Heninger, A.K., Poser, I., et al. (2010). A genome-scale DNA repair RNAi screen identifies SPG48 as a novel gene associated with hereditary spastic paraplegia. PLoS Biol. *8*.

Smith, G.A., Fearnley, G.W., Abdul-Zani, I., Wheatcroft, S.B., Tomlinson, D.C., Harrison, M.A., and Ponnambalam, S. (2016). VEGFR2 Trafficking, Signaling and Proteolysis is Regulated by the Ubiquitin Isopeptidase USP8. Traffic *17*, 53–65.

Sobreira, N., Schiettecatte, F., Valle, D., and Hamosh, A. (2015). GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. Hum. Mutat. *36*, 928–930.

Sogorb, M.A., Pamies, D., Estevan, C., Estévez, J., and Vilanova, E. (2016). Roles of NTE protein and encoding gene in development and neurodevelopmental toxicity. Chem. Biol. Interact. *259*, 352–357.

Sohaskey, M.L., Jiang, Y., Zhao, J.J., Mohr, A., Roemer, F., and Harland, R.M. (2010). Osteopotentia regulates osteoblast maturation, bone formation, and skeletal integrity in mice. J. Cell Biol. *189*, 511–525.

Souza, P.V.S., Pinto, W.B.V. de R., Batistella, G.N. de R., Bortholin, T., and Oliveira, A.S.B. (2016). Hereditary Spastic Paraplegia: Clinical and Genetic Hallmarks. Cerebellum *16*, 525–551.

Steinmuller, R., Lantigua-Cruz, A., Garcia-Garcia, R., Kostrzewa, M., Steinberger, D., and Muller, U. (1997). Evidence of a third locus in X-linked recessive spastic paraplegia [1]. Hum. Genet. *100*, 287–289.

Stevanin, G., Santorelli, F.M., Azzedine, H., Coutinho, P., Chomilier, J., Denora, P.S., Martin, E., Ouvrard-Hernandez, A.-M., Tessa, A., Bouslam, N., et al. (2007a). Mutations in SPG11, encoding spatacsin, are a major cause of spastic paraplegia with thin corpus callosum. Nat. Genet. *39*, 366–372.

Stevanin, G., Paternotte, C., Coutinho, P., Klebe, S., Elleuch, N., Loureiro, J.L., Denis, E., Cruz, V.T., Dürr, A., Prud'homme, J.F., et al. (2007b). A new locus for autosomal recessive spastic paraplegia (SPG32) on chromosome 14q12-q21. Neurology *68*, 1837–1840.

Stevanin, G., Azzedine, H., Denora, P., Boukhris, A., Tazir, M., Lossos, A., Rosa, A.L., Lerer, I., Hamri, A., Alegria, P., et al. (2008). 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.

Strümpell, A. (1880). Beiträge zur Pathologie des Rückenmarks. Arch. Psychiatr. Nervenkr. *10*, 676–717.

Subramony, S.H., Nguyen, T. V., Langford, L., Lin, X., Parent, A.D., and Zhang, J. (2009). Identification of a new form of autosomal dominant spastic paraplegia. Clin. Genet. *76*, 113–116.

Synofzik, M., and Schüle, R. (2017). Overcoming the divide between ataxias and spastic

paraplegias: Shared phenotypes, genes, and pathways. Mov. Disord. *0*, 1–14.

Tesson, C., Nawara, M., Salih, M.A.M., Rossignol, R., Zaki, M.S., Al Balwi, M., Schule, R., Mignot, C., Obre, E., Bouhouche, A., et al. (2012). Alteration of fatty-acid-metabolizing enzymes affects mitochondrial form and function in hereditary spastic paraplegia. Am. J. Hum. Genet. *91*, 1051–1064.

Tesson, C., Koht, J., and Stevanin, G. (2015). Delving into the complexity of hereditary spastic paraplegias: how unexpected phenotypes and inheritance modes are revolutionizing their nosology. Hum. Genet. *134*, 511–538.

Teuling, E., van Dis, V., Wulf, P.S., Haasdijk, E.D., Akhmanova, A., Hoogenraad, C.C., and Jaarsma, D. (2008). A novel mouse model with impaired dynein/dynactin function develops amyotrophic lateral sclerosis (ALS)-like features in motor neurons and improves lifespan in SOD1-ALS mice. Hum. Mol. Genet. *17*, 2849–2862.

Thomsen, B., Nissen, P.H., Agerholm, J.S., and Bendixen, C. (2010). Congenital bovine spinal dysmyelination is caused by a missense mutation in the SPAST gene. Neurogenetics *11*, 175–183.

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

Tiranti, V., Corona, P., Greco, M., Taanman, J.W., Carrara, F., Lamantea, E., Nijtmans, L., Uziel, G., and Zeviani, M. (2000). A novel frameshift mutation of the mtDNA COIII gene leads to impaired assembly of cytochrome c oxidase in a patient affected by Leigh-like syndrome. Hum. Mol. Genet. *9*, 2733–2742.

Travaglini, L., Bellacchio, E., Aiello, C., Pro, S., Bertini, E., and Nicita, F. (2017). Expanding the clinical phenotype of CAPN1-associated mutations: A new case with congenital-onset pure spastic paraplegia. J. Neurol. Sci. *378*, 210–212.

Tsang, H.T.H.H., Edwards, T.L., Wang, X., Connell, J.W., Davies, R.J., Durrington, H.J., O'Kane, C.J., Luzio, J.P., and Reid, E. (2009). The hereditary spastic paraplegia proteins NIPA1, spastin and spartin are inhibitors of mammalian BMP signalling. Hum. Mol. Genet. *18*, 3805–3821.

Tsaousidou, M.K., Ouahchi, K., Warner, T.T., Yang, Y., Simpson, M.A., Laing, N.G.G., Wilkinson, P.A., Madrid, R.E., Patel, H., Hentati, F., et al. (2008). Sequence Alterations within CYP7B1 Implicate Defective Cholesterol Homeostasis in Motor-Neuron Degeneration. Am. J. Hum. Genet. *82*, 510–515.

Valdmanis, P.N., Meijer, I.A., Reynolds, A., Lei, A., MacLeod, P., Schlesinger, D., Zatz, M., Reid, E., Dion, P.A., Drapeau, P., et al. (2007). Mutations in the KIAA0196 Gene at the SPG8 Locus Cause Hereditary Spastic Paraplegia. Am. J. Hum. Genet. *80*, 152–161.

Valente, E.M., Brancati, F., Caputo, V., Bertini, E., Patrono, C., Costanti, D., and Dallapiccola,

B. (2002). Novel locus for autosomal dominant pure hereditary spastic paraplegia (SPG19) maps to chromosome 9q33-q34. Ann. Neurol. *51*, 681–685.

Vazza, G., Zortea, M., Boaretto, F., Micaglio, G.F., Sartori, V., and Mostacciuolo, M.L. (2000). A New Locus for Autosomal Recessive Spastic Paraplegia Associated with Mental Retardation and Distal Motor Neuropathy, SPG14, Maps to Chromosome 3q27-q28. Am. J. Hum. Genet. *67*, 504–509.

Verny, C., Guegen, N., Desquiret, V., Chevrollier, A., Prundean, A., Dubas, F., Cassereau, J., Ferre, M., Amati-Bonneau, P., Bonneau, D., et al. (2011). Hereditary spastic paraplegia-like disorder due to a mitochondrial ATP6 gene point mutation. Mitochondrion *11*, 70–75.

Voeltz, G.K., Prinz, W.A., Shibata, Y., Rist, J.M., Rapoport, T.A., Hirling, H., Yamashita, K., Sasagawa, T., Yutsudo, M., Brosamle, C., et al. (2006). A class of membrane proteins shaping the tubular endoplasmic reticulum. Cell *124*, 573–586.

De Vos, K.J., and Hafezparast, M. (2017). Neurobiology of axonal transport defects in motor neuron diseases: Opportunities for translational research? Neurobiol. Dis. *105*, 283–299.

Wakil, S.M., Ramzan, K., Abuthuraya, R., Hagos, S., Al-Dossari, H., Al-Omar, R., Murad, H., Chedrawi, A., Al-Hassnan, Z.N., Finsterer, J., et al. (2014). Infantile-onset ascending hereditary spastic paraplegia with bulbar involvement due to the novel ALS2 mutation c.2761C>T. Gene *536*, 217–220.

Wang, D., Chan, C.-C.C., Cherry, S., and Hiesinger, P.R. (2013). Membrane trafficking in neuronal maintenance and degeneration. Cell. Mol. Life Sci. *70*, 2919–2934.

Wang, Y., Cameron, E.G., Li, J., Stiles, T.L., Kritzer, M.D., Lodhavia, R., Hertz, J., Nguyen, T., Kapiloff, M.S., and Goldberg, J.L. (2015). Muscle A-Kinase Anchoring Protein-α is an Injury-Specific Signaling Scaffold Required for Neurotrophic- and Cyclic Adenosine Monophosphate-Mediated Survival. EBioMedicine *2*, 1880–1887.

Wang, Y., Hersheson, J., Lopez, D., Hamad, M. Ben, Liu, Y., Pinto, V., Seinfeld, J., Wiethoff, S., Sun, J., Amouri, R., et al. (2016). Defects in the CAPN1 gene result in alterations in cerebellar development and in cerebellar ataxia in mice and humans. Cell Rep. *16*, 79–91.

Watanabe, F., Arnold, W.D., Hammer, R.E., Ghodsizadeh, O., Moti, H., Schumer, M., Hashmi, A., Hernandez, A., Sneh, A., Sahenk, Z., et al. (2013). Pathogenesis of Autosomal Dominant Hereditary Spastic Paraplegia (SPG6) Revealed by a Rat Model. J. Neuropathol. Exp. Neurol. *72*, 1016–1028.

Wedding, I.M., Koht, J., Tran, G.T., Misceo, D., Selmer, K.K., Holmgren, A., Frengen, E., Bindoff, L., Tallaksen, C.M.E., and Tzoulis, C. (2014). Spastic Paraplegia Type 7 Is Associated with Multiple Mitochondrial DNA Deletions.

Wilkinson, K.D., Lee, K., Deshpande, S., Duerksen-hughes, P., Boss, J.M., and Pohl, J. (1989). The Neuron-Specific Protein PGP 9.5 Is a Ubiquitin Carboxyl-Terminal Hydrolase. Science (80-.). *246*, 670–673.

Windpassinger, C., Auer-Grumbach, M., Irobi, J., Patel, H., Petek, E., Hörl, G., Malli, R., Reed, J. a, Dierick, I., Verpoorten, N., et al. (2004). Heterozygous missense mutations in BSCL2 are associated with distal hereditary motor neuropathy and Silver syndrome. Nat. Genet. *36*, 271–276.

Witte, K., Schuh, A.L., Hegermann, J., Sarkeshik, A., Mayers, J.R., Schwarze, K., Iii, J.R.Y., Eimer, S., Audhya, A., Yates Iii, J.R., et al. (2012). Mechanisms by which TFG functions in protein secretion and oncogenesis. Nat. Cell Biol. *13*, 550–558.

Wong, W., and Scott, J.D. (2004). AKAP signalling complexes: focal points in space and time. Nat. Rev. Mol. Cell Biol. *5*, 959–970.

Wong, W., Goehring, A.S., Kapiloff, M.S., Langeberg, L.K., and Scott, J.D. (2008). mAKAP compartmentalizes oxygen-dependent control of HIF-1alpha. Sci. Signal. *1*, ra18.

Wood, J.D., Landers, J.A., Bingley, M., McDermott, C.J., Thomas-McArthur, V., Gleadall, L.J., Shaw, P.J., and Cunliffe, V.T. (2006). The microtubule-severing protein Spastin is essential for axon outgrowth in the zebrafish embryo. Hum. Mol. Genet. *15*, 2763–2771.

Xiang, C., Wang, Y., Zhang, H., and Han, F. (2017). The role of endoplasmic reticulum stress in neurodegenerative disease. Apoptosis *22*, 1–26.

Xu, Q., and Huang, Y. (2006). Lipid metabolism in Alzheimer's and Parkinson's disease. Future Lipidol. *1*, 441–453.

Yadav, P.K., and Rajasekharan, R. (2016). DDHD domain-containing lipases: Targets for the treatment of rare diseases. J Rare Dis Res Treat. *1*, 34–38.

Yamamoto, Y., Yoshida, A., Miyazaki, N., Iwasaki, K., and Sakisaka, T. (2014). Arl6IP1 has the ability to shape the mammalian ER membrane in a reticulon-like fashion. Biochem. J. *458*, 69–79.

Yamanaka, T., and Nukina, N. (2018). ER Dynamics and Derangement in Neurological Diseases. Front. Neurosci. *12*, 91.

Yang, H., Sasaki, T., Minoshima, S., and Shimizu, N. (2007). Identification of three novel proteins (SGSM1, 2, 3) which modulate small G protein (RAP and RAB)-mediated signaling pathway. Genomics *90*, 249–260.

Yang, Y., Liu, W., Fang, Z., Shi, J., Che, F., He, C., Yao, L., Wang, E., and Wu, Y. (2015). A Newly Identified Missense Mutation in FARS2 Causes Autosomal-Recessive Spastic Paraplegia. Hum. Mutat.

Yasuda, T., Kaji, Y., Agatsuma, T., Niki, T., Arisawa, M., Shuto, S., Ariga, H., and Iguchi-Ariga, S.M.M. (2013). DJ-1 cooperates with PYCR1 in cell protection against oxidative stress. Biochem. Biophys. Res. Commun. *436*, 289–294.

Yoo, H., Son, D., Lee, Y.J., and Hong, K. (2016). Mouse JMJD4 is dispensable for embryogenesis. Mol. Reprod. Dev. *83*, 588–593.

Younes-Rapozo, V., Berendonk, J., Savignon, T., Manhães, A.C., and Barradas, P.C. (2006).

Thyroid hormone deficiency changes the distribution of oligodendrocyte/myelin markers during oligodendroglial differentiation in vitro. Int. J. Dev. Neurosci. *24*, 445–453.

Zeitlmann, L., Sirim, P., Kremmer, E., and Kolanus, W. (2001). Cloning of ACP33 as a novel intracellular ligand of CD4. J. Biol. Chem. *276*, 9123–9132.

Zhan, Z. xiong, Liao, X. xin, Du, J., Luo, Y. ying, Hu, Z. ting, Wang, J. ling, Yan, X. xiang, Zhang, J. guo, Dai, M. zhi, Zhang, P., et al. (2013). Exome sequencing released a case of X-linked adrenoleukodystrophy mimicking recessive hereditary spastic paraplegia. Eur. J. Med. Genet. *56*, 375–378.

Zhao, G., Hu, Z., Shen, L., Jiang, H., Ren, Z., Liu, X., Xia, K., Guo, P., Pan, Q., and Tang, B. (2008). A novel candidate locus on chromosome 11p14.1-p11.2 for autosomal dominant hereditary spastic paraplegia. Chin. Med. J. (Engl). *121*, 430–434.

Zhao, X., Alvarado, D., Rainier, S., Lemons, R., Hedera, P., Weber, C.H., Tukel, T., Apak, M., Heiman-Patterson, T., Ming, L., et al. (2001). Mutations in a newly identified GTPase gene cause autosomal dominant hereditary spastic paraplegia. Nat. Genet. *29*, 326–331.

Zhu, X., Liu, Y., Yin, Y., Shao, A., Zhang, B., Kim, S., and Zhou, J. (2009). MSC p43 required for axonal development in motor neurons. Proc. Natl. Acad. Sci. *106*, 15944–15949.

Zivony-Elboum, Y., Westbroek, W., Kfir, N., Savitzki, D., Shoval, Y., Bloom, A., Rod, R., Khayat, M., Gross, B., Samri, W., et al. (2012). A founder mutation in Vps37A causes autosomal recessive complex hereditary spastic paraparesis. J. Med. Genet. *49*, 462–472.

Zortea, M., Vettori, a, Trevisan, C.P., Bellini, S., Vazza, G., Armani, M., Simonati, a, and Mostacciuolo, M.L. (2002). Genetic mapping of a susceptibility locus for disc herniation and spastic paraplegia on 6q23.3-q24.1. J. Med. Genet. *39*, 387–390.

Züchner, S., Wang, G., Tran-Viet, K.-N., Nance, M. a, Gaskell, P.C., Vance, J.M., Ashley-Koch, A.E., Pericak-Vance, M.A., Zuchner, S., Tranviet, K., et al. (2006). Mutations in the novel mitochondrial protein REEP1 cause hereditary spastic paraplegia type 31. Am. J. Hum. Genet. *79*, 365–369.

Annex

Tables

Annex 1 - Table S1. Molecular pathways of HSP genes

Supplementary Articles

Annex 2 - Morais S, Bastos-Ferreira R, Sequeiros J, Alonso I: “Genomic mechanisms underlying *PARK2* large deletions identified in a cohort of patients with PD”. Neurol Genet. 2016 Jun; 2(3): e73

Annex 3 - Pereira MC, Morais S, Sequeiros J, Alonso I: “Large-Scale Functional RNAi screen in *C. elegans* Identifies TGF-ß and Notch Signaling Pathways as Modifiers of CACNA1A”. ASN Neuro. 2016 Mar 22;8(2)

Annex 4 - Matilla-Dueñas A, Corral-Juan M, Seuma A, Rolán D, Ispierto L, Morais S, Sequeiros J, Alonso I, Bertrán V, Serrano-Munuera C, Pintos-Morell G, Álvarez R, Sánchez I. “Chapter 25: Rare Neurodegenerative Diseases: Clinical and Genetic Update”. Adv Exp Med Biol. 2017;1031:443-496. doi: 10.1007/978-3-319-67144-4_25

Tables

Annex 1 - Table S1. Molecular pathways of HSP genes

Type	Chromosomal localization	Gene	Phenotype	Inheritance mode	Putative cellular function
SPG1	Xq28	***L1CAM***	C	XL	Cell adhesion and signaling
SPG2	Xq22.2	***PLP1***	P or C	XL	Myelination and Axonal preservation
SPG3A	14q22.1	***ATL1***	P or C	AD and AR	ER and Golgi morphogenesis, BMP signaling, Membrane traffic, Neurite outgrowth
SPG4	2p22.3	***SPAST***	P or C	AD	Microtubule dynamics, ER morphogenesis, Membrane and Endosomal traffic, BMP signaling, Cytokinesis
SPG5A	8q21.3	***CYP7B1***	P or C	AR	Cholesterol metabolism
SPG6	15q11.1	***NIPA1***	P or C	AD	Magnesium transport, BMP signaling, Endosomal traffic, Protein folding
SPG7	16q24.3	***SPG7***	P or C	AD and AR	Mitochondrial function
SPG8	8q24.13	***KIAA0196***	P or C	AD	Endosomal traffic, Cytoskeletal regulation, Protein folding
SPG9	10q24.1	***ALDH18A1***	P or C	AD and AR	Unknown (P5CS pathway)
SPG10	12q13	***KIF5A***	P or C	AD	Axonal transport
SPG11	15q21.1	***SPG11***	P or C	AR	Endosomal traffic, Autophagy
SPG12	19q13	***RTN2***	P	AD	ER morphogenesis, Membrane traffic
SPG13	2q33.1	***HSPD1***	P or C	AD	Mitochondrial function
SPG14	3q27-q28		P or C	AR	Unknown
SPG15	14q24.1	***ZFYVE26***	P or C	AR	Endosomal traffic, Autophagy, Cytokinesis
SPG16	Xq11.2		P or C	XL	Unknown
SPG17	11q13	***BSCL2***	C	AD	ER stress response, Lipid metabolism
SPG18	8p11.2	***ERLIN2***	C	AR	ER stress response, Lipid metabolism
SPG19	9q33-q34		P	AD	Unknown
SPG20	13q12.3	***SPG20***	C	AR	Endosomal traffic, BMP signaling, Lipid metabolism, Mitochondrial function, Microtubule dynamics, Cytokinesis
SPG21	15q22.31	***SPG21***	P or C	AR	Endosomal traffic and protein sorting
SPG22	Xq13.2	***SLC16A2***	P or C	XL	Membrane transporter, Axonal maturation, Myelination
SPG23	1q32.1	***DSTYK***	C	AR	Unknown
SPG24	13q14		P	AR	Unknown
SPG25	6q23-q24.1		C	AR	Unknown
SPG26	12q13.3	***B4GALNT1***	C	AR	Lipid metabolism, Myelination
SPG27	10q22.1-q24.1		P or C	AR	Unknown
SPG28	14q22.1	***DDHD1***	P or C	AR	Fatty-acid and/or phospholipid metabolism

Type	Chromosomal localization	Gene	Phenotype	Inheritance mode	Putative cellular function
SPG29	1p31.1-p21.1		C	AD	Unknown
SPG30	2q37	***KIF1A***	P or C	AD and AR	Axonal transport
SPG31	2p11.2	***REEP1***	P or C	AD	ER morphogenesis, ER-microtubule interaction, Mitochondrial function
SPG32	14q12-q21		C	AR	Unknown
SPG33	10q24.2	***ZFYVE27***	P	AD	ER morphogenesis, Membrane traffic
SPG34	Xq24-q25		P	XL	Unknown
SPG35	16q23.1	***FA2H***	P or C	AR	Myelination, Fatty acid metabolism
SPG36	12q23-24		C	AD	Unknown
SPG37	8p21.1-q13.3		P	AD	Unknown
SPG38	4p16-p15		C	AD	Unknown
SPG39	19p13	***PNPLA6***	C	AR	Phospholipid metabolism, Axonal integrity
SPG40	Reserved		P	AD	Unknown
SPG41	11p14.1-p11.2		P	AD	Unknown
SPG42	3q25.31	***SLC33A1***	P	AD	ER membrane transporter (Acetyl-CoA), BMP signaling, Autophagy
SPG43	19q12	***c19orf12***	C	AR	Lipid metabolism?, Autophagy?
SPG44	1q42.13	***GJC2***	C	AR	Myelination, Oligodendrocyte homeostasis
SPG45/65	10q24.33	***NT5C2***	P or C	AR	Purine/pyrimidine nucleotide metabolism
SPG46	9p13.3	***GBA2***	C	AR	Lipid metabolism
SPG47	1p13.2	***AP4B1***	C	AR	Vesicle formation and trafficking
SPG 48	7p22.1	***AP5Z1***	P or C	AR	Vesicle trafficking, Endosomal dynamics, DNA repair
SPG49	14q32.31	***TECPR2***	C	AR	Autophagy
SPG50	7q22.1	***AP4M1***	C	AR	Vesicle formation and trafficking
SPG51	15q21.2	***AP4E1***	C	AR	Vesicle formation and trafficking
SPG52	14q12	***AP4S1***	C	AR	Vesicle formation and trafficking
SPG53	8p22	***VPS37A***	C	AR	Vesicle trafficking
SPG54	8p11.23	***DDHD2***	C	AR	Lipid metabolism, Membrane traffic
SPG55	12q24.31	***C12orf65***	C	AR	Mitochondrial function
SPG56	4q25	***CYP2U1***	P or C	AR	Fatty acids metabolism
SPG57	3q12.2	***TFG***	C	AR	ER morphogenesis, Vesicle formation and trafficking
SPG58	17p13.2	***KIF1C***	P or C	AR	Retrograde vesicular transport from Golgi to ER
SPG59	15q21.2	***USP8***	C	AR	Endosomal morphology, Membrane traffic
SPG60	3p22.2	***WDR48***	C	AR	DNA damage repair, Lysosomal trafficking

Type	Chromosomal localization	Gene	Phenotype	Inheritance mode	Putative cellular function
SPG61	16p12.3	***ARL6IP1***	C	AR	ER morphogenesis, Protein transport, Membrane traffic?, Cell signaling?
SPG62	10q24.31	***ERLIN1***	P	AR	ER-associated degradation, Cholesterol homeostasis
SPG63	1p13.3	***AMPD2***	C	AR	Purine nucleotide metabolism
SPG64	10q24.1	***ENTPD1***	C	AR	Nucleotide metabolism
SPG66	5q32	***ARSI***	C	AR	Hormone biosynthesis, Cell signaling
SPG67	2q33.1	***PGAP1***	C	AR	GPI biosynthesis, ER-to-Golgi transport of GPI-anchor proteins
SPG68	11q13.1	***FLRT1***	C	AR	Cell adhesion and/or receptor signaling
SPG69	1q41	***RAB3GAP2***	C	AR	ER morphogenesis, Exocytosis of neurotransmitters and hormones
SPG70	12q13.3	***MARS***	C	AR	Protein biosynthesis?
SPG71	5p13.3	***ZFR***	P	AR	Unknown
SPG72	5q31.2	***REEP2***	P	AD and AR	ER morphogenesis
SPG73	19q13.33	***CPT1C***	P	AD	Lipid metabolism
SPG74	1q42.13	***IBA57***	C	AR	Mitochondrial function
SPG75	19q13.12	***MAG***	C	AR	Myelination
SPG76	11q13.1	***CAPN1***	C	AR	Axon maturation and maintenance
SPG77	6p25.1	***FARS2***	P	AR	Unknown (degeneration of Purkinje cells)
SPG78	1p36.13	***ATP13A2***	C	AR	Autophagy, Membrane trafficking
SPG79	4p13	***UCHL1***	C	AR	DNA damage response (deubiquitylating enzyme)
SPOAN	11q13.2	***KLC2***	C	AR	Axonal transport
SPG Un	9q22.31	***BICD2***	P or C	AD and AR	Axonal transport, Golgi dynamics
SPG Un	19p13.2	***DNM2***	P	AD	Endocytosis, Membrane trafficking
SPG Un	9p13.2	***EXOSC3***	C	AR	RNA exosome complex
SPG Un	2p25.1	***KIDINS220***	C	AD	Axon and dendritic maturation
SPG Un	6p21.1	***KLC4***	C	AR	Neuronal transport?
SPG Un	1q42.3	***LYST***	C	AR	Lysosomal traffic regulator
SPG Un	mitochondria	***MT-ATP6***	C	Maternally inherited SPG	Mitochondrial ATP production
SPG Un	1p36.22	***MTHFR***	C	AR	DNA and protein methylation, Lipid metabolism
SPG Un	19p13.3	***TUBB4A***	C	AD	Microtubule dynamics and stability

P – Pure; C – Complex; AD – Autosomal dominant; AR – Autosomal recessive; SPG Un – Unassigned SPG

Supplementary Articles

Annex 2 - Genomic mechanisms underlying *PARK2* large deletions identified in a cohort of patients with PD

Genomic mechanisms underlying *PARK2* large deletions identified in a cohort of patients with PD

OPEN

Sara Morais, MSc
Rita Bastos-Ferreira, MSc
Jorge Sequeiros, MD, PhD
Isabel Alonso, PhD

Correspondence to Dr. Alonso: ialonso@ibmc.up.pt

ABSTRACT

Objectives: To identify the genomic mechanisms that result in *PARK2* large gene deletions.

Methods: We conducted mutation screening using PCR amplification of *PARK2*-coding regions and exon-intron boundaries, followed by sequencing to evaluate a large series of 244 unrelated Portuguese patients with symptoms of Parkinson disease. For the detection of large gene rearrangements, we performed multiplex ligation-dependent probe amplification, followed by long-range PCR and sequencing to map deletion breakpoints.

Results: We identified biallelic pathogenic parkin mutations in 40 of the 244 patients. There were 18 different mutations, some of them novel. This study included mapping of 17 deletion breakpoints showing that nonhomologous end joining is the most common mechanism responsible for these gene rearrangements. None of these deletion breakpoints were previously described, and only one was present in 2 unrelated families, indicating that most of the deletions result from independent events.

Conclusions: The c.155delA mutation is highly prevalent in the Portuguese population (62.5% of the cases). Large deletions were present in 42.5% of the patients. We present the largest study on the molecular mechanisms that mediate *PARK2* deletions in a homogeneous population. ***Neurol Genet* 2016;2:e73; doi: 10.1212/NXG.0000000000000073**

GLOSSARY

AR-JP = autosomal recessive juvenile Parkinson disease; **CFS** = common fragile site; **MMEJ** = microhomology-mediated end joining; **MLPA** = multiplex ligation-dependent probe amplification; **NHEJ** = nonhomologous end joining; **NAHR** = nonallelic homologous recombination; **PD** = Parkinson disease.

Parkinson disease (PD) is the second most common neurodegenerative disorder, the etiology of which remains mostly unknown.[1] Most PD cases are sporadic, although the discovery of genes linked to familial forms has provided valuable insights into disease mechanisms. Among recessive forms, *PARK2* mutations are the most common cause of parkinsonism, being responsible for autosomal recessive juvenile Parkinson disease (AR-JP).[2] AR-JP is genetically heterogeneous and, apart from age at onset, is clinically indistinguishable from idiopathic PD, presenting with rigidity, bradykinesia, and resting tremor, usually before the age of 40 years.[3,4]

PARK2, located on chromosome 6q25.2–q27, encodes parkin, an E3 ubiquitin ligase. The loss of this ubiquitin ligase activity appears to be the mechanism responsible for the pathogenesis of AR-JP.[5] *PARK2* is composed of 12 exons surrounded by large intronic regions and spans approximately 1.38 Mb. Mutations have been identified across the entire gene and include all mutation types.[6] *PARK2* is the 17th largest gene of the human genome and is located within a large common fragile site (CFS), FRA6E,[7] a 3.6-Mb region of instability, susceptible to form gaps, breaks, and rearrangements when cells are exposed to certain conditions such as DNA replication inhibitors,[8–10] which may explain the large frequency of *PARK2* deletions.

From the UnIGENe (S.M., J.S., I.A.), Institute for Molecular and Cell Biology; i3S (S.M., R.B.-F., J.S., I.A.), Instituto de Investigação e Inovação em Saúde; CGPP (R.B.-F., J.S., I.A.), Institute for Molecular and Cell Biology; and Instituto de Ciências Biomédicas Abel Salazar (J.S., I.A.), University of Porto, Portugal.

Funding information and disclosures are provided at the end of the article. Go to Neurology.org/ng for full disclosure forms. The Article Processing Charge was paid by the authors.

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND), which permits downloading and sharing the work provided it is properly cited. The work cannot be changed in any way or used commercially.

Table 1 Overview of molecular and clinical information from 40 patients with a molecular diagnosis of autosomal recessive juvenile Parkinson disease

Patient	Age at onset, y	cDNA	Protein	Family history	Consanguinity
1	22	[c.1244C>A]+[c.172-52958_734+8943del]	[p.T415N]+[p.?]	Yes	No
2	28	[c.171+67708_734+58232delins28]+[c.171+67708_734+58232delins28]	[p.?]+[p.?]	No	No
3	38	[c.155delA]+[c.155delA]	[p.N52Mfs*29]+[p.N52Mfs*29]	No	Yes
4	25	[c.155delA]+[c.155delA]	[p.N52Mfs*29]+[p.N52Mfs*29]	Yes	No
5	17	[c.413-18966_871+72957delinsA]+[c.413-18966_871+72957delinsA]	[p.?]+[p.?]	No	Yes
6	32	[c.155delA]+[c.155delA]	[p.N52Mfs*29]+[p.N52Mfs*29]	Yes	Yes
7	30	[c.155delA]+[c.155delA]	[p.N52Mfs*29]+[p.N52Mfs*29]	No	No
8	27	[c.155delA]+[c.155delA]	[p.N52Mfs*29]+[p.N52Mfs*29]	No	No
9	13	[c.155delA]+[c.823C>T]	[p.N52Mfs*29]+[p.R275W]	Yes	No
10	31	[c.125G>C]+[c.8-51491_172-56018del]	[p.R42P]+[p.?]	No	No
11	35	[c.155delA]+[c.155delA]	[p.N52Mfs*29]+[p.N52Mfs*29]	No	No
12	25	[c.155delA] [c.155delA]	[p.N52Mfs*29]+[p.N52Mfs*29]	Yes	No
13	28	[c.155delA]+[c.172-11910_413-22473del]	[p.N52Mfs*29]+[p.?]	No	No
14	29	[c.534+457_734+41025del]+[c.534+457_734+41025del]	[p.?]+[p.?]	No	No
15	17	[c.823C>T]+[c.823C>T]	[p.R275W]+[p.R275W]	Yes	No
16	18	[c.1072_1073delinsA]+[c.1072_1073delinsA]	[p.L358Rfs*77]+[p.L358Rfs*77]	No	No
17	17	[c.171+67708_734+58232delins28]+[c.171+67708_734+58232delins28]	[p.?]+[p.?]	Yes	No
18	35	[c.155delA]+[c.155delA]	[p.N52Mfs*29]+[p.N52Mfs*29]	Yes	No
19	27	[c.155delA]+[c.155delA]	[p.N52Mfs*29]+[p.N52Mfs*29]	Yes	No
20	34	[c.1084-3859_1167+1618del]+[c.1084-3859_1167+1618del]	[p.?]+[p.?]	No	Yes
21	23	[c.155delA]+[c.735-21670_1083+48265delinsATG]	[p.N52Mfs*29]+[p.?]	Yes	No
22	51	[c.823C>T]+[c.413-16409_534+27042delinsATGCCTGTAA]	[p.R275W]+[p.?]	No	No
23	18	[c.155delA]+[c.155delA]	[p.N52Mfs*29]+[p.N52Mfs*29]	Yes	No
24	30	[c.155delA]+[c.8-23204_172-3140del]	[p.N52Mfs*29]+[p.?]	Yes	No
25	24	[c.413-3460_534+30928del]+[c.413-3460_534+30928del]	[p.?]+[p.?]	No	Yes
26	—	[c.155delA]+[c.155delA]	[p.N52Mfs*29]+[p.N52Mfs*29]	Yes	No
27	42	[c.155delA]+[c.155delA]	[p.N52Mfs*29]+[p.N52Mfs*29]	No	No
28	12	[c.155delA]+[c.155delA]	[p.N52Mfs*29]+[p.N52Mfs*29]	Yes	No
29	33	[c.125G>C]+[c.125G>C]	[p.R42P]+[p.R42P]	No	No
30	39	[c.155delA]+[c.412+25822_535-71707delinsTGA]	[p.N52Mfs*29]+[p.?]	No	No
31	50	[c.155delA]+[c.172-16570_734+51279del]	[p.N52Mfs*29]+[p.?]	Yes	No
32	13	[c.155delA]+[c.823C>T]	[p.N52Mfs*29]+[p.R275W]	Yes	No
33	38	[c.413-27055_534+20428del]+[c.413-27055_534+20428del]	[p.?]+[p.?]	No	No
34	52	[c.155delA]+[c.155delA]	[p.N52Mfs*29]+[p.N52Mfs*29]	Yes	Yes
35	39	[c.155delA]+[c.155delA]	[p.N52Mfs*29]+[p.N52Mfs*29]	Yes	No
36	38	[c.155delA]+[c.155delA]	[p.N52Mfs*29]+[p.N52Mfs*29]	Yes	No
37	20	[c.155delA]+[c.155delA]	[p.N52Mfs*29]+[p.N52Mfs*29]	No	No
38	55	[c.1097G>A]+[c.413-10504_534+408delins20]	[p.R366Q]+[p.?]	No	No
39	34	[c.1030delG]+[c.1286-3C>G]	[p.E344Sfs*91]+[p.?]	No	No
40	44	[c.155delA]+[c.412+2768_734+92226del]	[p.N52Mfs*29]+[p.?]	Yes	No

Table 2 Overview of 17 mapped deletions and responsible mechanisms

Patient	Deletion	State	Breakpoint upstream	Breakpoint downstream	Insertion	Mechanism	Deletion size, bp
1	3-6	Het	162.736.755	162.385.391	...	NHEJ	351.365
2	3-6	Hom	162.796.634	162.336.102	TTGATCCACCTGTCTCAGCAGGTGGAGA	NHEJ	460.533
5	4-7	Hom	162.641.250	162.133.847	A	NHEJ	507.404
10	2	Het	162.915.996	162.739.815	...	MMEJ	176.182
13	3	Het	162.695.707	162.644.757	AT	NHEJ	50.951
14	5-6	Hom	162.621.706	162.353.309	...	NHEJ	268.398
17	3-6	Hom	162.796.634	162.336.102	TTGATCCACCTGTCTCAGCAGGTGGAGA	NHEJ	460.533
20	10	Hom	161.811.768	161.806.208	...	*Alu*-mediated NAHR	5.561
21	7-9	Het	162.228.610	161.921.621	ATG	NHEJ	306.990
22	4	Het	162.638.694	162.595.120	CATGCCTGTAAT	NHEJ	43.575
24	2	Het	162.887.709	162.686.937	...	NHEJ	200.773
25	4	Hom	162.625.744	162.591.235	...	NHEJ	34.510
30	4	Het	162.657.735	162.546.913	TGA	NHEJ	110.823
31	3-6	Het	162.700.367	162.343.055	...	MMEJ	357.313
33	4	Hom	162.649.339	162.601.735	...	MMEJ	47.605
38	4	Het	162.632.788	162.621.753	TTACAGGTGAAATAAGATGCTT	NHEJ	11.036
40	4-6	Het	162.680.789	162.302.108	...	NHEJ	378.682

Abbreviations: Het = heterozygosity; Hom = homozygosity; MMEJ = microhomology-mediated end joining; NAHR = nonallelic homologous recombination; NHEJ = nonhomologous end joining.
Genomic coordinates are described according to the GRCh37 human reference assembly (hg19).

deletions could not be identified because of their small size or, in the case of the sequence inserted in patients 2 and 17, because of their frequency in the genome.

DISCUSSION PD, which affects almost 2% of the population above the age of 65 years, is the second most commonly occurring progressive neurodegenerative disorder after Alzheimer disease. Today, the lengthening lifespan and aging population has transformed neurodegenerative diseases into an economic and societal concern.[6] Mutations in the parkin coding gene are the predominant cause of monogenic forms of recessive PD with a frequency that varies across studies depending on the population, number of familial cases, age at onset, and consanguinity, among other factors.[11,12] The study of these mutations is therefore of upmost importance for understanding the pathogenesis of this disease.

Our *PARK2*-positive cohort showed a mean age at onset of 30 years (range, 12–55 years). This mutational study of the parkin coding gene revealed that 16.4% of these patients present with 2 mutations, thus confirming the PD clinical diagnosis. This is particularly interesting because we found several cases with parkin mutations with onset at more than 40 years of age. Also, the mean age at onset for patients in this study is lower than that reported previously (mean age 36 years) in a study considering only familial cases and with at least one of the patients with onset before the age of 54 years.[13] These findings show the importance of screening *PARK2* for mutations in patients with a family history of PD with later disease onset.

The mutations we identified are spread throughout the entire gene in almost every parkin domain. Of the 18 different mutations found, 3 were small insertions or deletions, 5 were point mutations, 9 were large deletions, and 1 was a splice-site mutation. Two were novel mutations predicted to be damaging to the protein: a small insertion (c.1030delG) and a splice-site mutation (c.1286-3C>G). It is interesting to note that we did not find duplications in our cohort.

Our results show that the c.155delA mutation is highly recurrent in the Portuguese population. The high frequency of this mutation in Portugal and Spain[14] probably indicates a founder effect. This mutation is also commonly found in European patients along with the c.823C>T mutation.[14–19] These 2 are the most common point mutations in our cohort and are located in exons 2 (ubiquitin-like domain) and 7 (RING1 domain). All missense mutations identified in our patients have been previously described.[6] Last, the splice-site mutation, c.1286-3C>G, is predicted to disrupt the acceptor splice site, resulting in the loss of the last exon.

Of the 9 identified deletions in our population, the most common were the exon 4 deletion (5 cases)

Table 3 Overview of repetitive elements found in breakpoint analysis

Patient	Deletion	Breakpoint upstream	Breakpoint downstream
1	3-6	Unique sequence	Unique sequence
2	3-6	Tigger 4A (DNA/TcMar-Tigger)	AluSx (SINE/Alu)
5	4-7	L1P4 (LINE/L1)	Mer21B (LTR/ERVL)
10	2	Unique sequence	Unique sequence
13	3	L1MC5 (LINE/L1)	L2C (LINE/L2)
14	5-6	L1ME3 (LINE/L1)	Unique sequence
17	3-6	Tigger 4A (DNA/TcMar-Tigger)	AluSx (SINE/Alu)
20	10	AluSx (SINE/Alu)	AluSx (SINE/Alu)
21	7-9	MLT1C (LTR/ERVL)	L1MC1 (LINE/L1)
22	4	AluSx (SINE/Alu)	Unique sequence
24	2	AluJb (SINE/Alu)	Unique sequence
25	4	Unique sequence	Unique sequence
30	4	AluSx (SINE/Alu)	Unique sequence
31	3-6	AluSx (SINE/Alu)	Unique sequence
33	4	Unique sequence	Unique sequence
38	4	Unique sequence	L1ME3 (LINE/L1)
40	4-6	Unique sequence	Unique sequence

and the exon 3–6 deletion (4 cases). The high frequency of these deletions is similar to the results of another study[17] that showed rearrangements occurring frequently in introns 2 through 4. Large deletions are responsible for juvenile PD in almost half of the studied patients, making MLPA an essential approach in molecular diagnosis, complementary to direct sequencing.

Large deletions account for approximately 50%–60% of the *PARK2* disease–causing mutations[8] and have been observed worldwide. This high frequency can be partially explained by the very large introns of *PARK2* that span 1.4 Mb,[8] and with introns 1, 2, 6, and 7 as large as 180 kb.[17] Another reason is the genomic context of *PARK2*. This gene is located within one of the 3 most frequently expressed CFS, namely FRA6E.[20] It is interesting to note that most of the rearrangements found in *PARK2* are located between exons 3 and 8, which comprise the FRA6E center, with exons 3 and 4 being among the most unstable regions and hotspots for deletion.[8,21] The cause of this instability is not completely understood, although several studies show that many gross chromosomal rearrangements that accumulate in solid tumors may have originated in these fragile sites because of errors during DNA replication that could be caused by intrinsic characteristics of CFS, such as their AT sequence content and high flexibility, scarcity of replication origins, and difficulty in the transcription of long genes present in these sites.[22]

To date, only a small number of *PARK2* deletion breakpoints have been mapped, probably because of the large gene size; despite the fact that the exonic deletions described here have been previously identified,[6] all the breakpoints present in these Portuguese patients are different from those previously described.[8,17,21,23–25]

In the 4 patients who had an exon 3–6 deletion, only 2 share the same breakpoints. In addition, all breakpoints were different in the 5 patients who had an exon 4 deletion. Thus, it seems that the same exon deletion is not necessarily a consequence of a similar rearrangement but may stem from independent events. This finding is in agreement with the ongoing hypothesis that rearrangements are independent and recurrent events.[24,26] It is also in accordance with the notion that rearrangements recurring at the same breakpoint are less frequent than nonrecurring rearrangements.[8] The high frequency of repetitive elements and instability in the *PARK2* locus favor the occurrence of recurrent deletion events, especially in the large *PARK2* introns.[17] It is curious that all of our patients with recurring deletions in homozygosity present the same breakpoints. In 3 patients, this could be explained by the presence of consanguinity. Two shared the same exon 3–6 deletion, which could indicate a recurrent event or a common founder mutation.

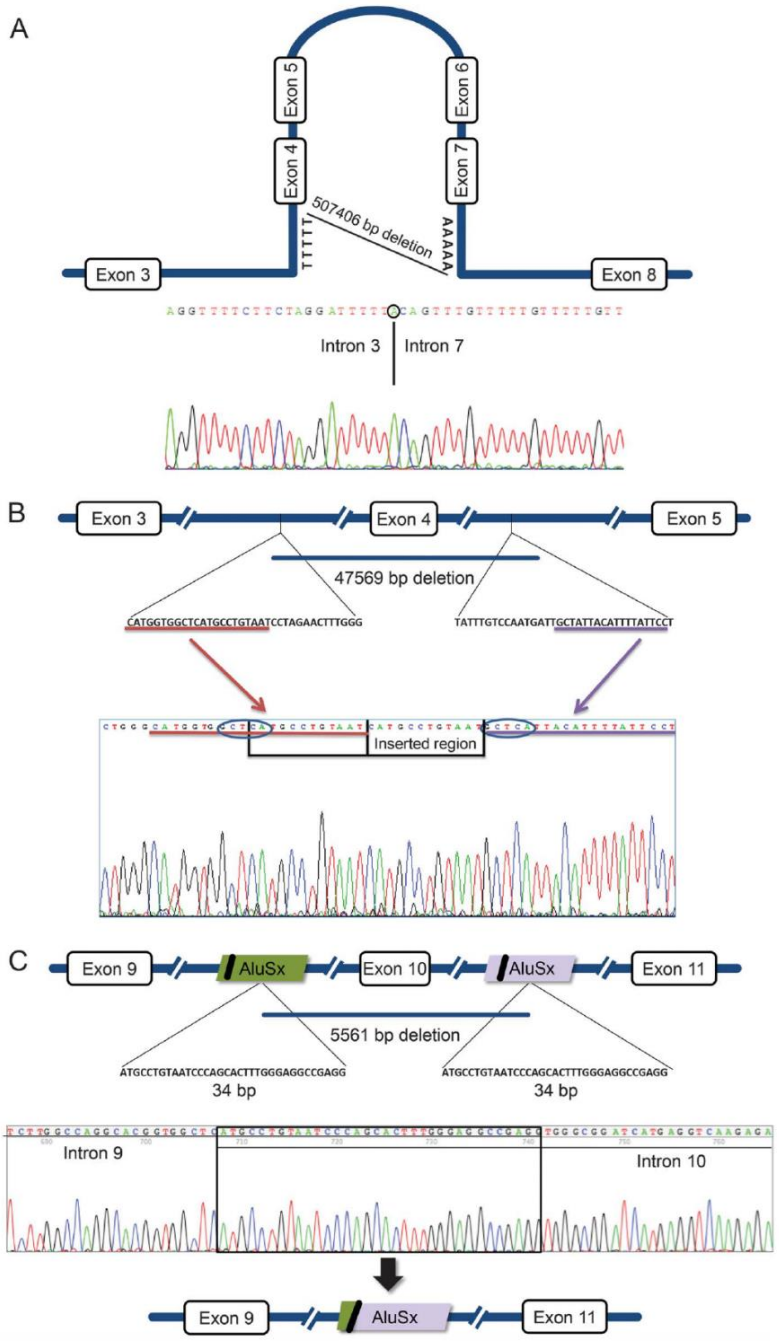
There are 3 major mechanisms responsible for genomic rearrangements: NAHR, NHEJ, and fork stalling and template switching. These 3 mechanisms account for the majority of genomic rearrangements in the human genome, and their distribution partially reflects the genomic architecture in the proximity of the breakpoint locus.[27]

Detailed analysis of the sequences flanking the breakpoints allowed us to explore causative mechanisms.

One mechanism implicated in *PARK2* deletions is NAHR, mediated by *Alu* elements. Our analysis showed that, although there are only 526 *Alu* elements in *PARK2*, which is below the mean density of the human genome (1 per 3 kb),[28] *Alu*-mediated NAHR seems to be the responsible mechanism for the exon 10 deletion described (figure, C) where the 2 junctions map to *Alu* elements of the same family (*Alu*Sx) in direct orientation. This mechanism has been previously described in *PARK2* deletions.[8]

Regions of microhomology often contribute to NHEJ, so it is common to find them at deletion breakpoints. NHEJ is the major pathway for restoring double-strand breaks in chromosomal DNA and is a highly flexible mechanism that creates distinct breakpoint junctions, resulting in either short microhomologies (usually 1–4 bp) or inserted sequences without homology as a result of the NHEJ editing process.[8] The presence of inserted sequences in the

Figure Schematic representation of 3 breakpoints found in our patients and the suspected mechanism for these deletions



(A) Nonhomologous end joining (NHEJ), mediated by the formation of a hairpin loop (deletion of exons 4–7; patient 5) and with the insertion of an adenine nucleotide. (B) NHEJ with the presence of a duplicated region and 2 motifs, complementary to immunoglobulin class switch repeats (deletion of exon 4; patient 22). (C) *Alu*-mediated nonallelic homologous recombination (NAHR) (deletion of exon 10; patient 20).

mapped deletions described here, namely a short duplication of the surrounding fragments and short sequences of unknown origin, is in agreement with features of the NHEJ mechanism. Also, this repair mechanism has already been related to parkin deletions in previous studies.[24,25] Nevertheless, a second mechanism is associated with the presence of microhomology. MMEJ is also a mechanism for double-strand break repair, which uses a sequence of 5–25 bp to align the broken ends; MMEJ is frequently associated with complex rearrangements. Although MMEJ is considered a secondary mechanism, used only when NHEJ and the other mechanisms fail, it has recently been reported that MMEJ can act even when NHEJ and homologous recombination are intact.[29] As 3 of the deletions reported here present a microhomology region of 5, 6, and 7 bp, we propose that this was the responsible mechanism in these cases (table 2). The other 5 cases present shorter microhomologies (table 2). Recently, 5 *PARK2* deletions were mapped in Polish patients showing NHEJ and fork stalling and template switching as the responsible mechanisms.[25] Here we present a higher number of deletions mapped, all different from previously described reports, and expanded the mechanisms responsible for these large gene rearrangements to include MMEJ and *Alu*-mediated NAHR. These 3 mechanisms, NHEJ, MMEJ, and *Alu*-mediated NAHR, seem to explain the identified deletions; NHEJ seems to be responsible for the majority of the deletions.

We describe 40 patients with autosomal recessive PD who have novel and previously reported *PARK2* mutations. We present the molecular characterization of 17 large *PARK2* deletions and identify NHEJ as the most frequently occurring mechanism responsible for deletions in this gene.

AUTHOR CONTRIBUTIONS

S. Morais: acquisition of data, analysis and interpretation, drafting and revision of manuscript. R. Bastos-Ferreira: acquisition of data, analysis and interpretation, and revision of manuscript. J. Sequeiros: revision of manuscript. I. Alonso: primary investigator, study conceptualization and design, analysis and interpretation, and revision of manuscript.

ACKNOWLEDGMENT

The authors thank all patients who participated in this study. They also thank all the physicians who referred patients to our center.

STUDY FUNDING

Supported by Fundação para a Ciência e Tecnologia cofunded by Fundo Europeu de Desenvolvimento Regional and COMPETE (Programa Operacional Factores de Competitividade), grant number PIC/IC/83232/2007.

DISCLOSURE

Ms. Morais and Ms. Bastos-Ferreira report no disclosures. Dr. Sequeiros has received travel/speaker honoraria from Fundación Jimenez Diaz and Pfizer; has served on the editorial boards of *Clinical Genetics* and the *Journal of Community Genetics*; is an employee of the University of Porto; and has received research support from Biomarin, Pfizer, Alexion, FCT (Fundação para a Ciência e Tecnologia), and the European Commission (FP6 Network of Excellence). Dr. Alonso reports no disclosures. Go to Neurology.org/ng for full disclosure forms.

Received November 15, 2015. Accepted in final form March 9, 2016.

REFERENCES

1. de Lau LM, Breteler MM. Epidemiology of Parkinson's disease. Lancet Neurol 2006;5:525–535.
2. Gasser T. Mendelian forms of Parkinson's disease. Biochim Biophys Acta 2009;1792:587–596.
3. von Coelln R, Dawson VL, Dawson TM. Parkin-associated Parkinson's disease. Cell Tissue Res 2004;318:175–184.
4. Moore DJ, West AB, Dawson VL, Dawson TM. Molecular pathophysiology of Parkinson's disease. Annu Rev Neurosci 2005;28:57–87.
5. Dawson TM, Dawson VL. The role of parkin in familial and sporadic Parkinson's disease. Mov Disord 2010;25 (suppl 1):S32–S39.
6. Nuytemans K, Theuns J, Cruts M, Van Broeckhoven C. Genetic etiology of Parkinson disease associated with mutations in the SNCA, PARK2, PINK1, PARK7, and LRRK2 genes: a mutation update. Hum Mutat 2010;31:763–780.
7. Smith DI, Zhu Y, McAvoy S, Kuhn R. Common fragile sites, extremely large genes, neural development and cancer. Cancer Lett 2006;232:48–57.
8. Mitsui J, Takahashi Y, Goto J, et al. Mechanisms of genomic instabilities underlying two common fragile-site-associated loci, PARK2 and DMD, in germ cell and cancer cell lines. Am J Hum Genet 2010;87:75–89.
9. Smith DI, McAvoy S, Zhu Y, Perez DS. Large common fragile site genes and cancer. Semin Cancer Biol 2007;17:31–41.
10. Denison SR, Callahan G, Becker NA, Phillips LA, Smith DI. Characterization of FRA6E and its potential role in autosomal recessive juvenile parkinsonism and ovarian cancer. Genes Chromosomes Cancer 2003;38:40–52.
11. Lucking CB, Durr A, Bonifati V, et al. Association between early-onset Parkinson's disease and mutations in the parkin gene. N Engl J Med 2000;342:1560–1567.
12. Kilarski LL, Pearson JP, Newsway V, et al. Systematic review and UK-based study of PARK2 (parkin), PINK1, PARK7 (DJ-1) and LRRK2 in early-onset Parkinson's disease. Mov Disord 2012;27:1522–1529.
13. Sun M, Latourelle JC, Wooten GF, et al. Influence of heterozygosity for parkin mutation on onset age in familial Parkinson disease: the GenePD study. Arch Neurol 2006;63:826–832.
14. Munoz E, Tolosa E, Pastor P, et al. Relative high frequency of the c.255delA parkin gene mutation in Spanish patients with autosomal recessive parkinsonism. J Neurol Neurosurg Psychiatry 2002;73:582–584.
15. Abbas N, Lucking CB, Ricard S, et al. A wide variety of mutations in the parkin gene are responsible for autosomal recessive parkinsonism in Europe. French Parkinson's Disease Genetics Study Group and the European Consortium on Genetic Susceptibility in Parkinson's Disease. Hum Mol Genet 1999;8:567–574.
16. Sironi F, Primignani P, Zini M, et al. Parkin analysis in early onset Parkinson's disease. Parkinsonism Relat Disord 2008;14:326–333.

17. Hedrich K, Eskelson C, Wilmot B, et al. Distribution, type, and origin of Parkin mutations: review and case studies. Mov Disord 2004;19:1146–1157.
18. Lohmann E, Periquet M, Bonifati V, et al. How much phenotypic variation can be attributed to parkin genotype? Ann Neurol 2003;54:176–185.
19. Bras J, Guerreiro R, Ribeiro M, et al. Analysis of Parkinson disease patients from Portugal for mutations in SNCA, PRKN, PINK1 and LRRK2. BMC Neurol 2008;8:1.
20. Gao G, Smith DI. Very large common fragile site genes and their potential role in cancer development. Cell Mol Life Sci 2014;71:4601–4615.
21. Clarimon J, Johnson J, Dogu O, et al. Defining the ends of Parkin exon 4 deletions in two different families with Parkinson's disease. Am J Med Genet B Neuropsychiatr Genet 2005;133B:120–123.
22. Franchitto A. Genome instability at common fragile sites: searching for the cause of their instability. Biomed Res Int 2013;2013:730714.
23. Elfferich P, Verleun-Mooijman MC, Maat-Kievit JA, et al. Breakpoint mapping of 13 large parkin deletions/duplications reveals an exon 4 deletion and an exon 7 duplication as founder mutations. Neurogenetics 2011;12:263–271.
24. Asakawa S, Hattori N, Shimizu A, et al. Analysis of eighteen deletion breakpoints in the parkin gene. Biochem Biophys Res Commun 2009;389:181–186.
25. Ambroziak W, Koziorowski D, Duszyc K, et al. Genomic instability in the PARK2 locus is associated with Parkinson's disease. J Appl Genet 2015;56:451–461.
26. Periquet M, Lucking C, Vaughan J, et al. Origin of the mutations in the parkin gene in Europe: exon rearrangements are independent recurrent events, whereas point mutations may result from founder effects. Am J Hum Genet 2001;68:617–626.
27. Gu W, Zhang F, Lupski JR. Mechanisms for human genomic rearrangements. Pathogenetics 2008;1:4.
28. Konkel MK, Batzer MA. A mobile threat to genome stability: the impact of non-LTR retrotransposons upon the human genome. Semin Cancer Biol 2010;20:211–221.
29. McVey M, Lee SE. MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. Trends Genet 2008;24:529–538.

Annex 3 - Large-Scale Functional RNAi screen in *C. elegans* Identifies TGF-ß and Notch Signaling Pathways as Modifiers of CACNA1A

American Society for Neurochemistry

Original Article

ASN Neuro
March-April 2016: 1–10
© The Author(s) 2016
DOI: 10.1177/1759091416637025
asn.sagepub.com
SAGE

Large-Scale Functional RNAi Screen in *C. elegans* Identifies TGF-β and Notch Signaling Pathways as Modifiers of *CACNA1A*

Maria da Conceição Pereira[1,2], Sara Morais[1,2], Jorge Sequeiros[1,2,3], and Isabel Alonso[1,2,3]

Abstract

Variants in *CACNA1A* that encodes the pore-forming α_1-subunit of human voltage-gated Cav2.1 (P/Q-type) Ca^{2+} channels cause several autosomal-dominant neurologic disorders, including familial hemiplegic migraine type 1, episodic ataxia type 2, and spinocerebellar ataxia type 6. To identify modifiers of incoordination in movement disorders, we performed a large-scale functional RNAi screen, using the *Caenorhabditis elegans* strain CB55, which carries a truncating mutation in the *unc-2* gene, the worm ortholog for the human *CACNA1A*. The screen was carried out by the feeding method in 96-well liquid culture format, using the ORFeome v1.1 feeding library, and time-lapse imaging of worms in liquid culture was used to assess changes in thrashing behavior. We looked for genes that, when silenced, either ameliorated the slow and uncoordinated phenotype of *unc-2*, or interacted to produce a more severe phenotype. Of the 350 putative hits from the primary screen, 37 genes consistently showed reproducible results. At least 75% of these are specifically expressed in the *C. elegans* neurons. Functional network analysis and gene ontology revealed overrepresentation of genes involved in development, growth, locomotion, signal transduction, and vesicle-mediated transport. We have expanded the functional network of genes involved in neurodegeneration leading to cerebellar ataxia related to *unc-2/CACNA1A*, further confirming the involvement of the transforming growth factor β pathway and adding a novel signaling cascade, the Notch pathway.

Keywords

RNAi screen, *unc-2*, enhancer/suppressor, disease modifiers

Received October 30, 2016; Received revised January 28, 2016; Accepted for publication January 31, 2016

Introduction

Voltage-gated calcium channels are involved in diverse biological processes of excitable cells, such as neurotransmitter release, excitation–contraction, regulation of gene expression, and neuronal migration by coupling electrical activity to intracellular Ca^{2+} signaling. These channels comprise an oligomeric assembly of different subunits that, when altered, were shown to be implicated in several cardiovascular and neurological diseases (Miller, 2001; Mathews et al., 2003).

The human *CACNA1A* gene encodes the α_1-subunit (pore-forming) of P/Q-type voltage-gated Ca^{2+} channels ($Ca_v2.1$). Mutations in this gene have been found to cause a spectrum of neuronal disorders with overlapping phenotypes: episodic ataxia type 2, familial hemiplegic migraine type 1, and spinocerebellar ataxia type 6 (Pietrobon, 2010; Rajakulendran et al., 2012).

CACNA1A is evolutionarily conserved, and its ortholog in *Caenorhabditis elegans* (*unc-2*) shares over 60% sequence identity with the human form. In the 1970s, Sydney Brenner's lab generated the CB55 strain, among others, by Ethyl methanesulfonate (EMS) mutagenesis

[1]UnIGENe, Institute for Molecular and Cell Biology (IBMC), Instituto de Investigação e Inovação em Saúde (i3S), University of Porto, Portugal
[2]Abel Salazar Institute for the Biomedical Sciences (ICBAS), University of Porto, Portugal
[3]CGPP, Institute for Molecular and Cell Biology (IBMC), Instituto de Investigação e Inovação em Saúde (i3S), University of Porto, Portugal

Corresponding Author:
Isabel Alonso, UnIGENe, Institute for Molecular and Cell Biology (IBMC), Instituto de Investigação e Inovação em Saúde (i3S), University of Porto, Portugal.
Email: ialonso@ibmc.up.pt

Creative Commons CC-BY: This article is distributed under the terms of the Creative Commons Attribution 3.0 License (http://www.creativecommons.org/licenses/by/3.0/) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).

from the N2 Bristol. The CB55 strain carries the *unc-2* canonical e55 mutant allele, that is, a base-pair substitution that introduces a premature stop codon, rendering the worm to an *unc-2* null phenotype (Mathews et al., 2003; Estevez et al., 2004). This truncating mutation is detected by the translation machinery and its mRNA targeted for nonsense-mediated decay. This culminates in the absence of *unc-2* expression or activity, effectively mimicking a knockout of the *unc-2* gene.

In *C. elegans*, *unc-2* is mainly expressed presynaptically, where its primary role is to mediate fast neurotransmission, although mosaic analysis has also suggested that it can be present both in the pre- and postsynaptic regions – neuronal and muscular functions (Schafer and Kenyon, 1995).

Strains carrying mutations in the *unc-2* gene are mainly characterized by slow and uncoordinated movement, slightly longer body length than wild-type (N2) animals, and constitutive egg-laying defects. The latter is due to their hypersensitivity to serotonin causing *unc-2* mutants to lay eggs with embryos at an abnormally early developmental stage and under inhibitory conditions for the N2 worms. At the molecular level, they show further defects in cholinergic and GABAergic synapses (Schafer and Kenyon, 1995; Mathews et al., 2003).

Overall, behavioral, pharmacological, and molecular phenotypes of the *unc-2* mutant strain closely reproduce the phenotypes associated with vertebrate P/Q-type channels (Rajakulendran et al., 2012). And, even though these genetic defects are rare, their study has already provided insight to common disorders of the central nervous system, such as migraine (Estevez et al., 2004, 2006) and epilepsy (Williams et al., 2004), and may continue further to do so.

To identify modifiers of the uncoordinated phenotype in *unc-2*-related channelopathies, we performed a large-scale functional RNAi screen, using the *C. elegans* strain CB55, which carries a truncating mutation in the *unc-2* gene, and looked for genes that, when silenced, either ameliorated the slow and uncoordinated phenotype of CB55 worms, or interacted to produce a more severe phenotype.

Materials and Methods

C. elegans Strains

Nematode strains were handled according to standard methods (Stiernagle, 2006). Strains were grown at 20°C, unless otherwise stated. N2 (Bristol) was used as the wild-type and the CB55 mutant strain, carrying a truncating mutation in the *unc-2* gene, presented the uncoordinated motor phenotype to be targeted by the modifier screen. All strains were provided by the Caenorhabditis Genetics Center, which is funded by the NIH Office of Research Infrastructure Programs (P40 OD010440).

RNAi Liquid Feeding

The screen was carried out in a 96-well liquid culture format using the *C. elegans* open reading frame (ORF)-RNAi feeding library (Source Bioscience), targeting approximately 11,000 genes, as detailed later.

Bacterial clones from the *C. elegans* ORFeome RNAi library v1.1 (Rual et al., 2004) were grown overnight in Lysogeny broth (LB) culture medium supplemented with ampicillin (100 µg/ml) and tetracycline (12.5 µg/ml) at 37°C and induced for 1 hr by adding Isopropyl β-D-1 thio-galactopyranoside (IPTG) to a final concentration of 4 mM, prior to incubation with the nematodes. Also, the previous day, hypochlorite egg extraction was performed to obtain synchronized L1 larvae (Stiernagle, 2006). The eggs were left to hatch overnight in the absence of food, at room temperature (20°C ± 5).

An average of 5 L1-stage worms in M9 Buffer (10µl) were distributed into 96-wells containing 40 µl of RNAi feeding bacteria, suspended in Nematode growth medium (NGM) media supplemented with ampicillin, tetracycline, and IPTG (Lehner et al., 2006). Feeding plates were afterward incubated at 20°C for 7 days with shaking. The H115 strain carrying the empty vector pL4440 and the *hil-5* ORF (previously reported to have no effect in the worm phenotype; Fraser et al., 2000; Rual et al., 2004; Sonnichsen et al., 2005) were used to control bacterial feeding and the standalone induction of the RNAi machinery. The *lpin-1* ORF, producing a severe developmental phenotype (Rual et al., 2004; Sonnichsen et al., 2005), was used to control RNAi induction efficiency in every 96-well plate.

Time-lapse imaging of worms in liquid culture was used to assess changes in thrashing behavior of the *C. elegans* strains. Every well was recorded for 30 s at 10 Hz, using an automated recording stage coupled to a charge-coupled device-equipped stereoscope. A single thrash was defined as a complete change in the direction of the body down the midline.

Raw imaging data were analyzed with open source wrMTrck plugin for Image J, and the thrashing results were loaded on CellHTS2 for further analysis (Pelz et al., 2010). One important aspect of data acquisition and analysis regards full automation, which allowed the reduction of random errors and bias introduced by the experimenter.

A primary screen was carried out on the CB55 mutant strain with duplicates and, after analysis, RNAi hits were considered when separated from the mean at least two standard deviations. Genes fulfilling these criteria were filtered for severe phenotypes, such as Emb and Ste, already reported in WormBase (WS235; Chen et al., 2005); these were excluded from further analysis. The remaining clones were taken into a second and third assay rounds, using the approach described earlier, with the exception that both the N2 and CB55 strains were

used in parallel and experiments were run in triplicates. This confirmation assays had the purpose to fulfil two goals: first, to confirm reproducibility of the effect observed in the primary screen; and, second, to identify genetic interactions specific of the CB55 mutant.

All final clone hits were directly sequenced to verify their identity using the M13 universal primer sequence.

Functional Gene Networks

The top 40 genes related to the ones identified in the genetic screen were plotted using Cytoscape v3.3.1 (Shannon et al., 2003) and GeneMANIA App (Montojo et al., 2010), which allowed for network generation, according to Gene ontology (GO) term enrichment analysis and gene function prediction (only terms with Q value $\leq 10^{-4}$ were considered).

Results and Discussion

The aim of our study was to gather information on the genes and pathways that may act as modifiers in diseases associated to the *CACNA1A* gene. To accomplish this, we performed a genetic screen in the uncoordinated CB55 *C. elegans* strain, followed by ontology analysis to identify the main biological processes that acted as suppressors or enhancers of this mutant phenotype.

Screen Design

Because the *CACNA1A* gene is physiologically related to an expanding spectrum of human neurological disorders that implicate transient or progressive motor deficits, we performed a genetic screening for suppressors and enhancers of the motor phenotype in a model system. In *C. elegans*, the *CACNA1A* ortholog, *unc-2*, is highly homologous to its human counterpart and, when mutated, causes uncoordinated body movement, among other features that greatly reproduce the phenotype associated with human disease.

We took the CB55 strain and performed a large-scale functional RNAi screen in liquid culture, using the ORFeome library (Lamesch et al., 2004), targeting approximately 11,000 *C. elegans* genes.

Due to the mutant's hyperactive egg-laying behavior, the average number of worms distributed per well had to be reduced to the minimum number that guarantied no empty wells.

Another concern, partially related to the egg-laying behavior of this strain, is the development rate. While for the N2 wild-type strain one can observe and score progeny after 4 days of feeding (at 20°C), the CB55 worms need more time to grow and reach adulthood (probably due to slower pharyngeal pumping and, therefore, poor feeding), and also, as the eggs are prematurely laid, they take significantly longer to hatch. All these reasons account for the 7 days of feeding we found to be required prior to progeny imaging and scoring of thrashing behavior. Figure 1 shows an overview of the screen workflow and conditions.

unc-2 *Enhancer/Suppressor Large-Scale RNAi Screen*

Primary screen. A primary screen was carried out on the CB55 mutant strain with duplicates, and, after image acquisition and

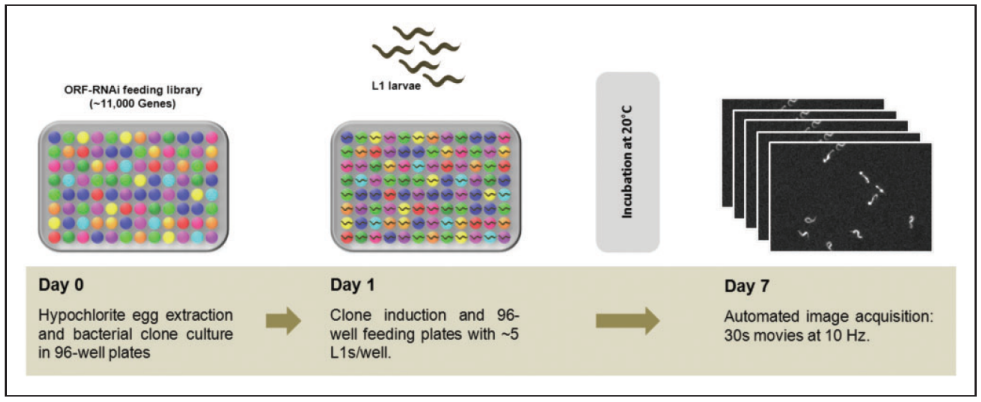


Figure 1. RNAi screen workflow. On the first day (Day 0), a 96-well plate was inoculated with the RNAi bacterial clones, and the CB55 eggs are extracted by bleaching gravid worms. The larvae then hatched overnight, and in the absence of food, they arrest at the L1-stage. Next day (Day 1), after induction of RNAi expression, each bacterial clone is fed to approximately five synchronous worms, which during the period of incubation grow and lay eggs. On Day 7, the progeny has hatched, and each well is imaged for 30 s to record their movement. ORF = open reading frame.

thrash analysis, the raw data were further explored through CellHTS2. Because the worm nervous system is known to be somewhat refractory to systemic RNAi, only the progeny was considered during the analysis, as they had been exposed to partial gene depletion since embryogenesis.

Data analysis revealed a mean score distribution around 1.005 ± 0.438, meaning that approximately 95% of the RNAi tested had no effect over the CB55 phenotype. We considered as preliminary RNAi hits, those genes with scores at least two standard deviations from the mean. Figure 2 shows a graphical representation of the CellHTS2 score results series, highlighting those outside the mean $\pm 2SD$. The primary screen yielded 676 putative RNAi hits. Among these, however, were putative enhancers, producing severe phenotypes that have been previously reported in WormBase (WS235; Chen et al., 2005) for the N2 wild-type strain. Several RNAi already reported to produce embryonic lethality (Emb) and adult sterility (Ste) were excluded from further analysis, which reduced the number of potential hits to 350, for further confirmation.

Validation screens. The remaining bacterial clones from the primary screen (350) were taken into two consecutive assays using the same approach with the exception that both the N2 and CB55 strains were used in parallel and experiments were run in triplicates. This allowed us to confirm reproducibility of the phenotype observed in the primary screen, and, by introducing the N2 wild-type strain, we were able to discriminate between genetic interactions specific of the *unc-2* mutant and those common to both strains, as well as discordant effects (Table 1). Overall, bacterial clones targeting 37 genes consistently showed reproducible results. Groups II to VI include discordant RNAi effects between N2 and CB55, while group VII comprises a list of genes that when downregulated increase motor ability in both strains, having a more generic effect. Furthermore, taking into account the available expression data in WormBase (WS242; Chen et al., 2005), at least 75% of the genes identified are known to be expressed in the *C. elegans* neurons.

Reduced Notch Signaling Aggravates CB55 Incoordination

The *unc-2* genetic interaction screen allowed the identification of several genes that, when downregulated, have an effect on the CB55 strain's uncoordinated movement. Further functional network analysis, using

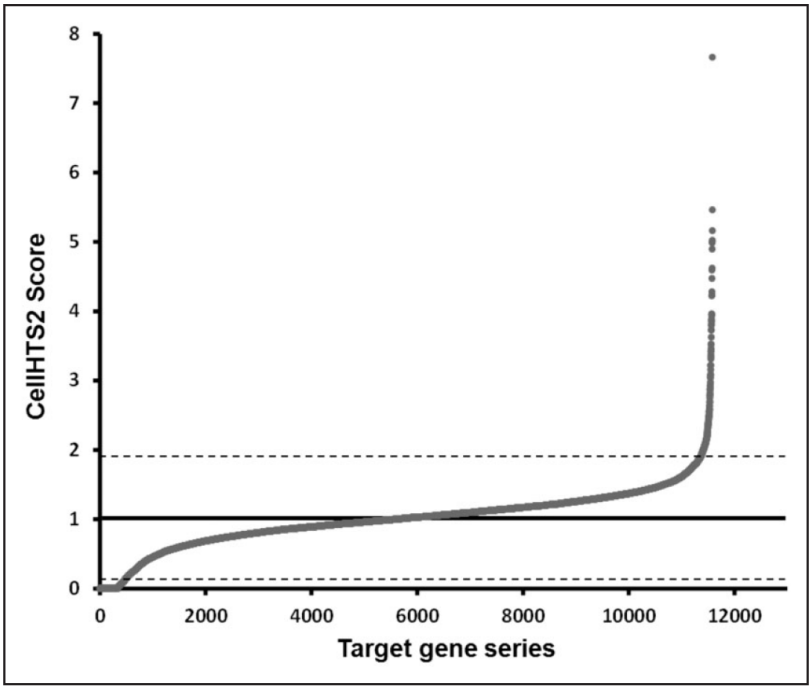


Figure 2. Graphical representation of the CellHTS2 score results series of the primary screen. Data analysis using CellHTS2 revealed a mean score distribution around 1.005 ± 0.438, that is, $\geq 95\%$ of the RNAi tested had no effect over the CB55 phenotype. Genes were considered preliminary RNAi hits, when scoring at least two standard deviations away from the mean (highlighted by the dashed line). At this stage, the screen yielded 676 putative RNAi hits.

Table 1. Genetic Enhancers and Suppressors of the *unc-2* Motor Phenotype.

		Thrashing rate (score)			
Group	Gene	N2 (3.3)	CB55 (0.7)	Expression	Protein
I	*ncbp-1*	− (0.0)	− (0.6)	N/A	NCBP-1 – Nuclear Cap-Binding Protein, orthologous to human NCBP1
II	*vps-45*	= (4.0)	− (0.0)	Nervous system	VPS-45 – Related to yeast Vacuolar Protein Sorting factor
	M57.2	= (3.7)	− (0.0)	N/A	N/A
	snr-5	= (4.1)	− (0.5)	Nervous system	SNR-5 – Small Nuclear Ribonucleoprotein
III	*cdc-6*	+ (5.1)	− (0.2)	Seam cell	CDC-6 – Cell Division Cycle related
	xpf-1	+ (3.4)	− (0.4)	Gonad	XPF-1 – (Xeroderma Pigmentosum group F) DNA repair gene homolog
	npax-1	+ (4.2)	− (0.2)	Head neurons	NPAX-1 – N-terminal PAX (PAI domain only) protein
	sago-2	+ (4.9)	− (0.2)	N/A	SAGO-2 – Synthetic secondary siRNA-deficient ArGOnaute mutant
	top-1	+ (4.6)	− (0.1)	Nervous system	TOP-1 – TOPoisomerase
	klp-3	+ (4.2)	− (0.0)	Marginal cell	KLP-3 – Kinesin-Like Protein
	dbl-1	+ (4.1)	− (0.1)	Nervous system	DBL-1 – DPP/BMP-Like
	sec-8	+ (6.2)	− (0.0)	N/A	SEC-8 – yeast SEC homolog
	Y47G6A.19	+ (4.4)	− (0.1)	N/A	Y47G6A.19 – orthologous to human carboxypeptidase A and B
	ppfr-4	+ (5.6)	− (0.1)	N/A	PPFR-4 – Protein Phosphatase Four Regulatory subunit
	aph-2	+ (4.4)	− (0.2)	Oocytes	APH-2 – ortholog of nicastrin
IV	*unc-97*	− (0.0)	= (0.6)	Nervous system	UNC-97 – LIM domain-containing protein of the PINCH family
V	*egl-43*	− (1.4)	+ (1.4)	Head neurons	EGL-43 – Zinc finger protein
	ceh-21	− (1.1)	+ (1.8)	Nervous system	CEH-21 –*Caenorhabditis elegans* Homeobox
VI	*daf-7*	= (3.5)	+ (0.9)	ASI neurons	TGF-β
	tbx-38	= (3.2)	+ (1.3)	Head neurons	TBX-38 – T box transcription factor
	prx-5	= (3.1)	+ (1.4)	Nervous system	PRX-5 – ortholog of the human receptor for type I peroxisomal targeting signal protein
	K10C3.5	= (3.1)	+ (2.2)	N/A	K10C3.5 – ortholog of budding yeast Ria1p/Efl1p and human EFTUD1
	unc-43	= (3.2)	+ (0.9)	Nervous system	UNC-43 – ortholog of type II calcium/calmodulin-dependent protein kinase (CaMKII)
	pat-12	= (3.3)	+ (101)	Reproductive system, hypodermis, rectal epithelium	PAT-12 – Paralysed Arrest at Twofold
	sedl-1	= (3.4)	+ (0.9)	N/A	SEDL-1 – human SEDL (spondyloepiphyseal dysplasia tarda) related
	Y54E10A.10	= (3.1)	+ (1.7)	N/A	N/A
	unc-51	= (3.1)	+ (1.0)	Nervous system	UNC-51 – serine/threonine protein kinase orthologous to *Saccharomyces cerevisiae* autophagy protein Atg1p and the vertebrate ULK proteins
	sma-2	= (3.4)	+ (2.2)	Head neurons	SMA-2 – orthologous to Smad1

(continued)

Table 1. Continued

Group	Gene	Thrashing rate (score) N2 (3.3)	CB55 (0.7)	Expression	Protein
VII	*flp-4*	+ (4.6)	+ (1.3)	Nerve ring	FLP-4 – FMRF-Like Peptide
	nhr-74	+ (4.9)	+ (0.9)	Seam cell	NHR-74 – Nuclear Hormone Receptor family
	sma-6	+ (4.9)	+ (1.0)	ASI neurons	SMA-6 – serine/threonine protein kinase orthologous to type I TGF-β receptor
	daf-14	+ (4.0)	+ (0.9)	Nerve ring, AVJ, RIV	DAF-14 – orthologous to Smad2
	nkcc-1	+ (7.1)	+ (0.9)	N/A	nkcc-1 – Na-K-Cl Cotransporter homolog
	R107.5	+ (5.5)	+ (.09)	N/A	N/A
	ntr-1	+ (5.0)	+ (1.5)	ASEL, ADF, BDU, ASH, PQR	NTR-1 – NemaTocin Receptor, orthologous to the human vasopressin receptor type 2
	pqn-85	+ (4.1)	+ (0.9)	N/A	PQN-85 – Prion-like-(Q/N-rich)- domain-bearing protein, orthologous to human NIPBL
	Y51H1A.3	+ (4.1)	+ (0.9)	N/A	Y51H1A.3 – orthologous to human NDUFB8

Note. Motor function: (+) improved, (=) no change, and (−) worsened; N/A = information not available; TGF = Transforming growth factor β: DPP/BMP=Bone morphogenic protein.

Cytoscape (Figure 3), allowed us to reduce the complexity of the obtained data and pinpoint individual and interconnected pathways with a major impact on phenotype. Among these is the Notch signaling pathway. We have found that *aph-2* silencing results in an aggravation of *unc-2* uncoordinated phenotype. This gene is the worm ortholog of nicastrin, one of the proteins that constitute γ-secretase, a multimeric protease composed of presenilin, nicastrin, Aph1, and Pen2 (Crump et al., 2013); γ-secretase cleaves the amyloid precursor protein to release the Aβ peptide known to be involved in Alzheimer's disease pathogenesis. In addition, this protease also cleaves Notch, a key step of the Notch signaling pathway, which plays an essential role not only in cell fate decisions during development but also in the adult brain, including maintenance and differentiation of neuronal stem cells, structural and synaptic plasticity, as well as neuronal survival (Ables et al., 2011). To date, no enhancers for the *unc-2* phenotype are known. From the genes belonging to the Notch pathway, *lin-12* (Notch) and *sel-12* (PEN-1/2) are not available in the RNAi library, *pen-2* (PEN2) knockout levels showed no effect on CB55, *aph-2* (Nicastrin) consistently worsened the phenotype, and *aph-1* (APH1), going back to the initial screen, scored on top of the cutoff for worsening the phenotype and thus did not carry on for further analysis, nevertheless showing a clear trend for half of the gamma secretase complex.

Table 1 highlights a phenomenon very common in many neurological disorders—context dependency—as discordant effects on N2 and CB55 stains arise from knockdown with the same RNAi. The phenotypic outcome is dependent on the presence or absence of a functional *unc-2* channel.

Our results suggest that downregulation of nicastrin, and perhaps APH1, may result in decreased Notch cleavage and consequent impairment of Notch-dependent transcriptional activation, aggravating the *unc-2* motor phenotype. Consistent with our findings, the important role of this pathway in cerebellar ataxia is further strengthened by the fact that other Notch signaling components (RBP-J/Su(H)) have been identified as modulators of pathogenesis in a *Drosophila* SCA17 model (Ren et al., 2011). Moreover, ataxin-1, the protein mutated in SCA1, and its related protein BOAT1 have also been found to be components of the Notch signaling pathway (Tong et al., 2011).

TGFβ and CaMKII Signaling Downregulation Partially Rescues the unc-2 *Phenotype*

The gene network plotted in Figure 3 clearly evidences the importance of the TFG- β pathway for the *unc-2* genetic background. We found that downregulation of several genes involved in this signaling pathway partially

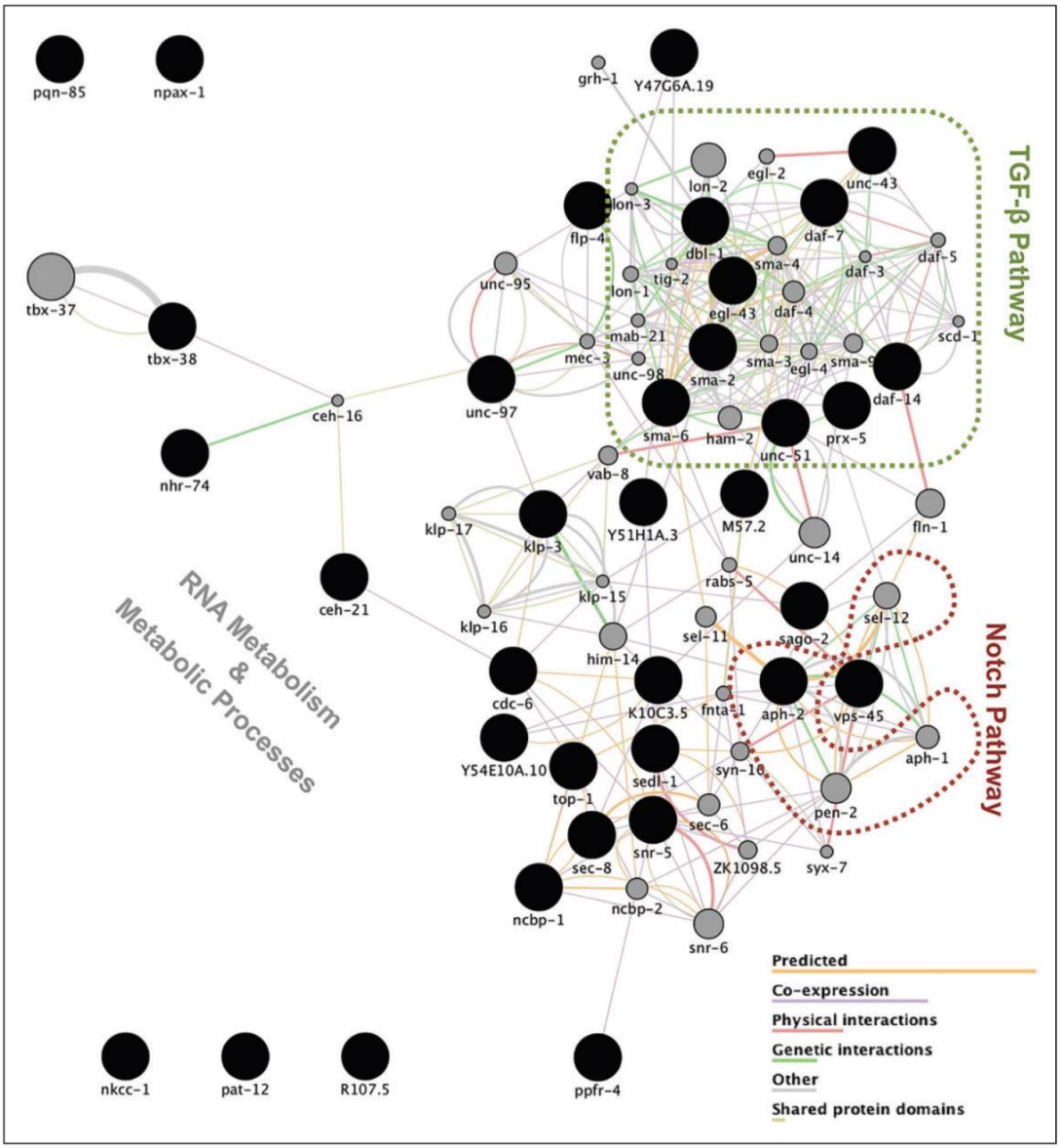


Figure 3. Gene network of enhancer and suppressor genes of the *unc-2* phenotype. The top 40 genes related to those 37 identified in the genetic screen were plotted using Cytoscape v3.3.1 and GeneMANIA App, which allowed for the network generation according to GO term enrichment analysis and gene function prediction. TGF-β and Notch signaling cascades were the main pathways to be identified by the genetic screen as involved in modulating *unc-2* lethargy. Black-colored circles depict direct hits from the screen, while gray-colored ones represent the 40 genes most related to hits, sized according to their weight in the network. The significance of the involvement of other genes relating to RNA metabolism and basic cellular metabolic processes, in light of the available data, remains elusive.
TGF-β = transforming growth factor β.

rescue the *unc-2* uncoordinated phenotype. Among these are *daf-7*, which encodes the worm ortholog for the transforming growth factor β (TGF-β) superfamily; *sma-2*, a receptor-regulated Smad protein (R-Smad), orthologous to Smad1; *sma-6*, the ortholog for the type I TGF-β receptor; and *daf-14* the worm ortholog for Smad2. In fact, partial knockdown of *sma-6* and *daf-14* not only improved CB55 motor function but also had a

positive impact in the motor ability of the wild-type strain, as evidenced in Table 1.

The TGF-β is a pleiotropic cytokine that regulates a diverse range of cellular processes, such as proliferation, differentiation, migration, and apoptosis. In humans, the various TGF-β isoforms are expressed in neurons, and glial cells and its receptors are found throughout the central nervous system (Shi and Massague, 2003). TGF-β response is mediated by a high-affinity transmembrane receptor serine/threonine kinase complex, consisting of the TGF-β type I (TRI) and type II (TRII) receptors. Ligand interaction with a TRII dimer recruits and activates TRI, which in turn phosphorylates the receptor-regulated Smad proteins (Smad2 and 3), which are then translocated into the nucleus where they regulate transcription of their target genes (Patterson and Padgett, 2000; Shi and Massague, 2003). This process is well conserved in *C. elegans* (Patterson and Padgett, 2000).

Our results reinforce the previous association of the TGF-β pathway to the lethargic phenotype of the CB55 strain (Estevez et al., 2004). In their study, Estevez et al. crossed null strains for various components of the worm TGF-β pathway to find that ablation, one-by-one, of most of these genes was able to partially rescue *unc-2* lethargy to different degrees. In our screen, only some of these genes were shown to change the phenotype, what can be readily accounted by partial depletion versus null alleles; however, they were fully sufficient to identify the entire pathway through network analysis.

The TGF-β family of cytokines comprises two subfamilies: the TGF-β/Activin/Nodal and the bone morphogenetic protein/growth and differentiation factor/Müllerian-inhibiting substance subfamily (Shi and Massague, 2003). Interestingly, both in a previous study (Estevez et al., 2004) and in this screen, the *dbl-1* gene, which encodes a bone morphogenetic protein-like protein that binds the TGF-β receptor, has either shown no effect, or in our case, to enhance motor deficits, raising the possibility that, in the *unc-2* background, the signaling pathway dissociates based on the ligand subfamily initiating the cascade to produce a different response.

Unc-2-dependent Ca^{2+} influx has long been hypothesized to be linked to *unc-43*/CaMKII regulation, impacting gene expression symmetry in the neurons (Troemel et al., 1999), neuronal migration (Tam et al., 2000), and tyrosine hydroxylase (*tph-1*) upregulation (Estevez et al., 2004). Here, we show that partial depletion of *unc-43* has no effect over N2 coordination; however, in the *unc-2* null background, it partially rescues lethargy, confirming its role in the *unc-2* signaling cascade. Although the combination of these two null alleles has failed to clearly establish this connection, because the outcome was enhanced lethargy, one could argue that the null allele might not be the best suitable model for a protein of critical importance in many cellular pathways, such as *unc-43*. In fact, it has been shown that different levels of downregulation are more sensitive in terms of detecting genetic interactions among essential genes (Fievet et al., 2013).

Apart from *dbl-1*, all other members of the TGF-β cascade align perfectly with the expected partial suppression of the phenotype (*daf-7*, *daf-14*, *sma-2*, and *sma-6*). Looking back into the primary screen, *daf-8* and *sma-3* scored right on the cut-off line for improving the phenotype, and although no further testing was performed, it only supports the overall trend: partial suppression of the *unc-2* phenotype. Even *unc-43* that in previous studies was shown to have unexpected enhanced lethargy, partially improved CB55's motility in our study, most likely due to partial knockdown instead of full knockout.

The only gene identified that is directly downstream of *unc-2* signaling is *unc-43* (CaMKII; Figure 4). The other genes, including those belonging to the previously associated TGF-β pathway, are in parallel signaling cascades to *unc-2*. Figure 4 depicts other major cell signaling mechanisms that crosstalk with insulin and TGF-β (previously established to interact with the *unc-2* cascade; Estevez et al., 2006; Guo and Wang, 2009; Andersson et al., 2011) such as the Wnt and TKR pathways that may very well be a source of more *unc-2* modifiers. Our library covered only half the Notch pathway genes, and a small minority of genes involved in Wnt and TRK signaling, thus the question on their possible involvement remains.

We have expanded the functional network of genes related to *unc-2* functional role in disease, from the previously reported serotonergic and insulin signaling (Estevez et al., 2006), further confirming the involvement of the TGF-β pathway and adding the Notch pathway as a novel signaling cascade.

The fact that all *unc-2* suppressor genes found to date only partially rescue the phenotype argues in favor of a multifactorial network of modifiers that greatly reflect the phenotypic variability of patients carrying *CACNA1A* mutations, with presentation of symptoms in diverse combinations.

Conclusions

Our results provide new insight into the genes and pathways that may contribute to neuronal dysfunction and death in calcium channel-related cerebellar ataxia, with an emphasis in specific cellular pathways, particularly TGF-β and Notch signaling. Different approaches have been used to search for modifiers and deregulated pathways in neurodegenerative disorders. We have combined a large-scale genetic screen in *C. elegans*, with functional network analysis, to identify genes and pathways

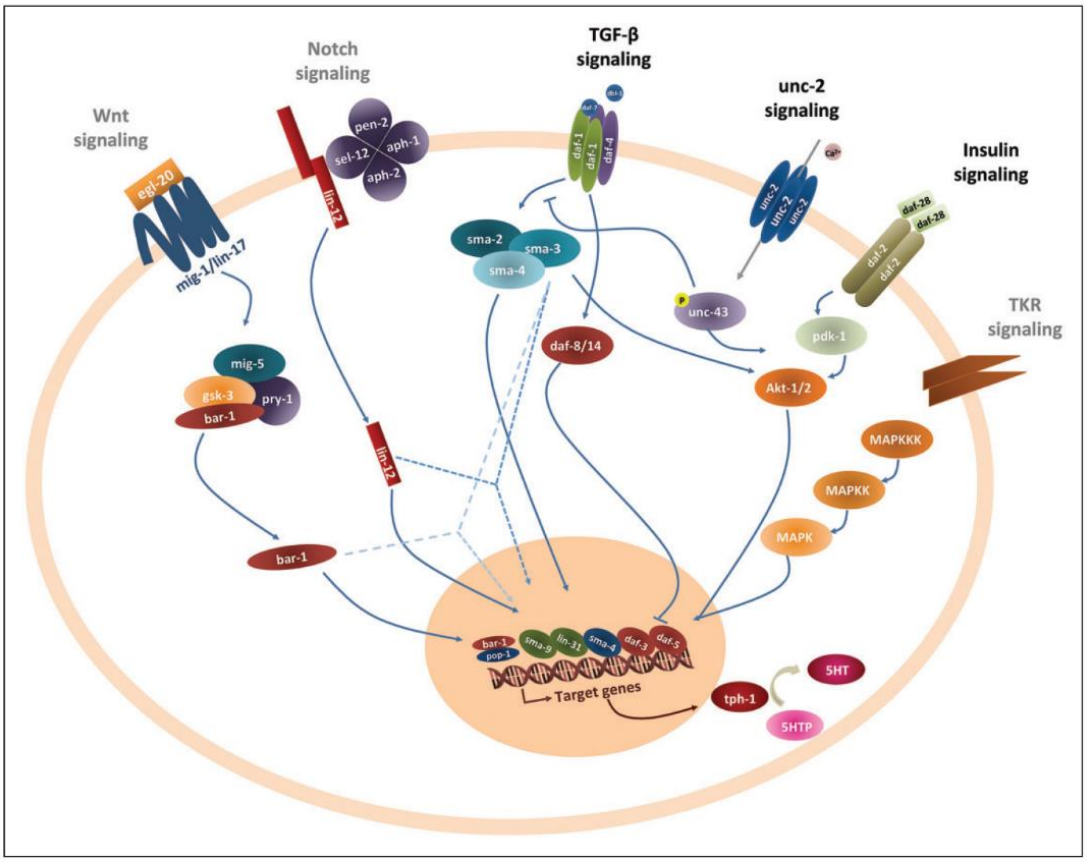


Figure 4. Crosstalk between *unc*-2 signaling and other cell signaling pathways. Key intracellular mediators of the *unc*-2, insulin, TGF-β, Notch, TRK, and Wnt pathways are depicted. Pathways identified by black heading are already known modifiers of the *unc*-2 phenotype, the remaining, colored gray, depict parallel pathways that, at a given point down the cascade, interact with the already established *unc*-2, insulin, or TGF-β pathways (dashed lines). The unc-2 null allele (CB55) modifier screens support that the majority of genetic modifiers (enhancers or suppressors) are more likely in different parallel pathways like insulin and TGF-β. The figure shows other major cell signaling mechanisms that crosstalk with insulin and TGF-β and may be a source of more *unc*-2 modifiers.
TGF-β = transforming growth factor β.

important for cerebellar ataxia-related neurodegeneration that may constitute candidate targets for therapeutical interventions.

Author Contributions

M. C. P. performed most of the experiments, discussed and interpreted data, and drafted the document. S.M. performed the experiments and revised the document. J. S. revised the document. I. A. designed the experiments, discussed and interpreted data, and revised the document.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: this work was supported by Fundação para a Ciência e Tecnologia (PTDC/SAU-GMG/100913/2008), cofunded by FEDER (Fundo Europeu de Desenvolvimento Regional) and COMPETE (Programa Operacional Factores de Competitividade), and by Financiamento Plurianual de Unidades de Investigação (FCT).

References

Ables, J. L., Breunig, J. J., Eisch, A. J., & Rakic, P. (2011). Not(ch) just development: Notch signalling in the adult brain. *Nature Reviews. Neuroscience*, *12*, 269–283.

Andersson, E. R., Sandberg, R., & Lendahl, U. (2011). Notch signaling: Simplicity in design, versatility in function. *Development*, *138*, 3593–3612.

Chen, N., Harris, T. W., Antoshechkin, I., Bastiani, C., Bieri, T., Blasiar, D., . . . Stein, L. D. (2005). WormBase: A comprehensive data resource for Caenorhabditis biology and genomics. *Nucleic Acids Research*, *33*, D383–D389.

Crump, C. J., Johnson, D. S., & Li, Y. M. (2013). Development and mechanism of gamma-secretase modulators for Alzheimer's disease. *Biochemistry*, *52*, 3197–3216.

Estevez, A. O., Cowie, R. H., Gardner, K. L., & Estevez, M. (2006). Both insulin and calcium channel signaling are required for developmental regulation of serotonin synthesis in the chemosensory ADF neurons of Caenorhabditis elegans. *Developmental Biology*, *298*, 32–44.

Estevez, M., Estevez, A. O., Cowie, R. H., & Gardner, K. L. (2004). The voltage-gated calcium channel UNC-2 is involved in stress-mediated regulation of tryptophan hydroxylase. *Journal of Neurochemistry*, *88*, 102–113.

Fievet, B. T., Rodriguez, J., Naganathan, S., Lee, C., Zeiser, E., Ishidate, T., . . . Ahringer, J. (2013). Systematic genetic interaction screens uncover cell polarity regulators and functional redundancy. *Nature Cell Biology*, *15*, 103–112.

Fraser, A. G., Kamath, R. S., Zipperlen, P., Martinez-Campos, M., Sohrmann, M., & Ahringer, J. (2000). Functional genomic analysis of *C. elegans* chromosome I by systematic RNA interference. *Nature*, *408*, 325–330.

Guo, X., & Wang, X. F. (2009). Signaling cross-talk between TGF-beta/BMP and other pathways. *Cell Research*, *19*, 71–88.

Lamesch, P., Milstein, S., Hao, T., Rosenberg, J., Li, N., Sequerra, R., . . . Vidal, M. (2004). C. elegans ORFeome version 3.1: Increasing the coverage of ORFeome resources with improved gene predictions. *Genome Research*, *14*, 2064–2069.

Lehner, B., Tischler, J., & Fraser, A. G. (2006). RNAi screens in Caenorhabditis elegans in a 96-well liquid format and their application to the systematic identification of genetic interactions. *Nature Protocols*, *1*, 1617–1620.

Mathews, E. A., Garcia, E., Santi, C. M., Mullen, G. P., Thacker, C., Moerman, D. G., & Snutch, T. P. (2003). Critical residues of the Caenorhabditis elegans unc-2 voltage-gated calcium channel that affect behavioral and physiological properties. *The Journal of Neuroscience*, *23*, 6537–6545.

Miller, R. J. (2001). Rocking and rolling with Ca2+ channels. *Trends in Neurosciences*, *24*, 445–449.

Montojo, J., Zuberi, K., Rodriguez, H., Kazi, F., Wright, G., Donaldson, S. L., . . . Bader, G. D. (2010). GeneMANIA Cytoscape plugin: Fast gene function predictions on the desktop. *Bioinformatics*, *26*, 2927–2928.

Patterson, G. I., & Padgett, R. W. (2000). TGF beta-related pathways. *Roles in Caenorhabditis elegans development.* Trends in Genetics, *16*, 27–33.

Pelz, O., Gilsdorf, M., & Boutros, M. (2010). web cellHTS2: A web-application for the analysis of high-throughput screening data. *BMC Bioinformatics*, *11*, 185.

Pietrobon, D. (2010). CaV2.1 channelopathies. *Pflügers Archiv: European Journal of Physiology*, *460*, 375–393.

Rajakulendran, S., Kaski, D., & Hanna, M. G. (2012). Neuronal P/Q-type calcium channel dysfunction in inherited disorders of the CNS. *Nature Reviews. Neurology*, *8*, 86–96.

Ren, J., Jegga, A. G., Zhang, M., Deng, J., Liu, J., Gordon, C. B., . . . Ma, J. (2011). A Drosophila model of the neurodegenerative disease SCA17 reveals a role of RBP-J/Su(H). *in modulating the pathological outcome*. Human Molecular Genetics, *20*, 3424–3436.

Rual, J. F., Ceron, J., Koreth, J., Hao, T., Nicot, A. S., Hirozane-Kishikawa, T., . . . Vidal, M. (2004). Toward improving Caenorhabditis elegans phenome mapping with an ORFeome-based RNAi library. *Genome Research*, *14*, 2162–2168.

Schafer, W. R., & Kenyon, C. J. (1995). A calcium-channel homologue required for adaptation to dopamine and serotonin in Caenorhabditis elegans. *Nature*, *375*, 73–78.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., . . . Ideker, T. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Research*, *13*, 2498–2504.

Shi, Y., & Massague, J. (2003). Mechanisms of TGF-beta signaling from cell membrane to the nucleus. *Cell*, *113*, 685–700.

Sonnichsen, B., Koski, L. B., Walsh, A., Marschall, P., Neumann, B., Brehm, M., . . . Echeverri, C. J. (2005). Full-genome RNAi profiling of early embryogenesis in Caenorhabditis elegans. *Nature*, *434*, 462–469.

Stiernagle, T. (2006). *Maintenance of C. elegans* WormBook, pp. 1–11.

Tam, T., Mathews, E., Snutch, T. P., & Schafer, W. R. (2000). Voltage-gated calcium channels direct neuronal migration in Caenorhabditis elegans. *Developmental Biology*, *226*, 104–117.

Tong, X., Gui, H., Jin, F., Heck, B. W., Lin, P., Ma, J., . . . Tsai, C. C. (2011). Ataxin-1 and Brother of ataxin-1 are components of the Notch signalling pathway. *EMBO Reports*, *12*, 428–435.

Troemel, E. R., Sagasti, A., & Bargmann, C. I. (1999). Lateral signaling mediated by axon contact and calcium entry regulates asymmetric odorant receptor expression in *C. elegans*. *Cell*, *99*, 387–398.

Williams, S. N., Locke, C. J., Braden, A. L., Caldwell, K. A., & Caldwell, G. A. (2004). Epileptic-like convulsions associated with LIS-1 in the cytoskeletal control of neurotransmitter signaling in Caenorhabditis elegans. *Human Molecular Genetics*, *13*, 2043–2059.

Annex 4 – Chapter 25.3 Hereditary Spastic Paraplegias

Chapter 25
Rare Neurodegenerative Diseases: Clinical and Genetic Update

Antoni Matilla-Dueñas, Marc Corral-Juan, Agustí Rodríguez-Palmero Seuma, Dolores Vilas, Lourdes Ispierto, Sara Morais, Jorge Sequeiros, Isabel Alonso, Víctor Volpini, Carmen Serrano-Munuera, Guillem Pintos-Morell, Ramiro Álvarez, and Ivelisse Sánchez

A. Matilla-Dueñas (*) • M. Corral-Juan
Functional and Translational Neurogenetics Unit, Department of Neurosciences, Health Sciences Research Institute Germans Trias-IGTP, Can Ruti Campus, Ctra de Can Ruti, Camí de les Escoles s/n, 08916 Badalona, Barcelona, Spain

A. Rodríguez- Palmero Seuma • G. Pintos-Morell
Department of Pediatrics, University Hospital Germans Trias i Pujol (HUGTP) and Health Sciences Research Institute, Can Ruti Campus, Ctra. de Canyet s/n, 08916 Badalona, Barcelona, Spain

D. Vilas • L. Ispierto • R. Álvarez
Neurodegenerative Diseases Unit, Neurology Service and Neurosciences Department, University Hospital Germans Trias i Pujol (HUGTP), Ctra. de Canyet s/n, Can Ruti Campus, 08916 Badalona, Barcelona, Spain

S. Morais • J. Sequeiros • I. Alonso
IBMC – Institute for Molecular and Cell Biology, i3S – Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Porto, Portugal

V. Volpini
Bellvitge Biomedical Research Institute (IDIBELL), L'Hospitalet, Barcelona, Spain

C. Serrano-Munuera
Department of Internal Medicine, Hospital St. Joan de Déu, Martorell, Spain
Manresa University, Martorell, Barcelona, Spain

I. Sánchez (*)
Functional Biology and Experimental Therapeutics Laboratory, Functional and Translational Neurogenetics Unit, Department of Neurosciences, Health Sciences Research Institute Germans Trias-IGTP, Can Ruti Campus, Ctra de Can Ruti, Camí de les Escoles s/n, 08916 Badalona, Barcelona, Spain

© Springer International Publishing AG 2017
M. Posada de la Paz et al. (eds.), Rare Diseases Epidemiology: Update and Overview, Advances in Experimental Medicine and Biology 1031, https://doi.org/10.1007/978-3-319-67144-4_25

25.3 Hereditary Spastic Paraplegias

The hereditary spastic paraplegias (HSPs) were first identified by Seeligmüller, Strümpell and Lorrain as an autosomal dominant disease, characterized by progressive spasticity and weakness of the lower limbs, with moderate loss of vibratory sense and bladder dysfunction. At that time, the neuropathological hallmark of the disease was also described as the degeneration of the longest spinal pathways, corticospinal tracts, and medial dorsal columns. The classification of HSPs is difficult and, throughout the years, several proposals have been made based on phenotype, mode of inheritance, and mutated gene (SPGs). All modes of hereditary transmission are found: autosomal dominant (AD), autosomal recessive (AR), X-linked (XL), and mitochondrial inheritance. Clinically, the HSPs have been subdivided into pure and complex forms, according to the presence or absence of other neurological and extra-neurological features.

25.3.1 Clinical Manifestations

The initial symptoms in HSP patients include a feeling of stiffness, muscle cramps, inability to walk rapidly and frequent falls. In early-onset cases the disease is often expressed as delayed gait acquisition. Age-at-onset is highly variable, particularly for pure forms, ranging from the first year of life to the 8th decade, tending to be later in autosomal dominant forms and earlier in recessive ones. At disease onset, spasticity is usually noticeable only while walking. Over time, especially in complex forms, pyramidal signs may affect the upper limbs, though many patients show only tendon hyperreflexia that may include a brisk jaw reflex; weakness or spasticity of the upper limbs is rare, particularly in pure forms. In some patients with complex forms, dysarthria and dysphagia may present as a pseudobulbar state. Other manifestations include cognitive impairment (mental retardation or deterioration), epilepsy, optic atrophy, amyotrophies, neuropathy (usually axonal), ataxia and dystonia [15].

Until now, 89 loci and 75 genes have been identified: 20 autosomal dominant, 57 autosomal recessive, five X-linked, one with mitochondrial inheritance, and 6 with both dominant and recessive transmission (Table 25.4). A recent study has identified HSP mutations in genes associated with Parkinson (ATP13A2/SPG78), neuronal ceroid lipofuscinosis (TPP1), and the hereditary motor and sensory neuropathy (DNMT1), highlighting the genetic, in addition to the clinical, heterogeneity of spastic paraplegia [17].

Table 25.4 Genetics of spastic paraplegias

Name	Locus	OMIM	Gene	Mode of inheritance	Average onset (range in years)	Distinguishing clinical features
SPG1	Xq28	303350	*L1CAM*	X-linked	Infancy	Ataxia, mental retardation, hydrocephalus.
SPG2	Xq22.2	312920	*PLP1*	X-linked	Infancy or childhood	Pure or with cerebellar dysfunction, hypotonia, dementia, seizures, mental retardation.
SPG3A	14q22.1	182600	*ATL1*	AD and AR	Before 10 (1–68	Pure or rarely axonal neuropathy
SPG4	2p22.3	182601	*SPAST*	AD	30 (1–80)	Pure or neuropathy, cognitive impairment, agitation.
SPG5A	8q21.3	270800	*CYP7B1*	AR	Variable (1–47)	Pure or cerebellar signs, nystagmus, cognitive impairment.
SPG6	15q11.1	600363	*NIPA1*	AD	16.5 (9–35)	Pure or rarely neuropathy.
SPG7	16q24.3	602783	*SPG7*	AD and AR	30 (25–42)	Pure or optic atrophy, supranuclear palsy, ataxic gait, pyramidal signs.
SPG8	8q24.13	603563	*KIAA0196*	AD	Adult onset (18–60)	Pure or atrophy of shins.
SPG9	10q24.1	601162	*ALDH18A1*	AD and AR	Variable (1–30)	Pure or amyotrophy, cataracts, neuropathy, gastroesophageal reflux.
SPG10	12q13	604187	*KIF5A*	AD	Variable (2–51)	Pure or neuropathy, amyotrophy.
SPG11	15q21.1	610844	*SPG11*	AR	Early variability (1–30)	Pure or amyotrophy, neuropathy, cognitive decline, cerebellar signs.
SPG12	19q13	604805	*RTN2*	AD	6.9 (5–22)	Pure.
SPG13	2q33.1	605280	*HSPD1*	AD	39 (17–68)	Pure or pyramidal signs.
SPG14	3q27-q28	605229	Unknown	AR	30	Pure or mental retardation, distal motor neuropathy.
SPG15	14q24.1	270700	*ZFYVE26*	AR	Infancy or childhood (5–19)	Pure or distal amyotrophy, cerebellar signs, ataxia, cognitive deterioration, axonal neuropathy.
SPG16	Xq11.2	300266	Unknown	X-linked	Infancy	Pure or quadriplegia, motor aphasia, mental retardation, pyramidal signs.
SPG17	11q13	270685	*BSCL2*	AD	Variable (2–60)	Neuropathy, distal limb muscle atrophy and weakness.
SPG18	8p11.2	611225	*ERLIN2*	AR	Infancy (1–6)	Mental retardation, contractures, primary lateral sclerosis.
SPG19	9q33-q34	607152	Unknown	AD	47 (36–55)	Pure.
SPG20	13q12.3	275900	*SPG20*	AR	Early childhood	Troyer syndrome, distal amyotrophy, cerebellar signs.
SPG21	15q22.31	248900	*SPG21*	AR	Adulthood	Mast syndrome, frontotemporal atrophy, pyramidal signs.
SPG22	Xq13.2	300523	*SLC16A2*	X-linked	Infancy	Mental retardation, dystonia, ataxia, quadriplegia
SPG23	1q32.1	270750	*DSTYK*	AR	Early childhood	Hyperpigmentation, cognitive impairment, peripheral neuropathy.
SPG24	13q14	607584	Unknown	AR	Early onset	Pure.
SPG25	6q23-q24.1	608220	Unknown	AR	Adulthood (30–46)	Sensory or motor neuropathy, pyramidal signs secondary to spinal cord compression.
SPG26	12q13.3	609195	*B4GALNT1*	AR	Childhood (2–19)	Dystonia, ataxia, distal amyotrophy, mental retardation.
SPG27	10q22.1-q24.1	609041	Unknown	AR	Adulthood (25–45)	Pure or sensorimotor polyneuropathy, dysarthria, hyperactive bladder.
SPG28	14q22.1	609340	*DDHD1*	AR	Childhood (7–15)	Pure or pyramidal signs, distal sensory impairment in lower limbs.
SPG29	1p31.1-p21.1	609727	Unknown	AD	Infancy	Neonatal hyperbilirubinemia, auditory neuropathy, hiatal hernia.
SPG30	2q37.3	610357	*KIF1A*	AD/AR	Variable (10–39)	Pure axonal neuropathy or cerebellar signs.
SPG31	2p11.2	610250	*REEP1*	AD	30 (2–45)	Pure distal sensory loss, amyotrophy.
SPG32	14q12-q21	611252	Unknown	AR	Childhood	Mild mental retardation, cerebellar atrophy.

Name	Locus	OMIM	Gene	Mode of inheritance	Average onset (range in years)	Distinguishing clinical features
SPG33	10q24.2	610244	*ZFYVE27*	AD	Adulthood (42–50)	Pure.
SPG34	Xq24-q25	300750	Unknown	X-linked	Adulthood (10–25)	Pure.
SPG35	16q23.1	612319	*FA2H*	AR	(3–11)	Dystonia, optic atrophy, ataxia, cognitive decline, seizures.
SPG36	12q23-q24	613096	Unknown	AD	24 (14–33)	Demyelinating motor and sensory neuropathy.
SPG37	8p21.1-q13.3	611945	Unknown	AD	32 (8–60)	Pure.
SPG38	4p16-p15	612335	Unknown	AD	Young adulthood (16–19)	Pure.
SPG39	19p13.2	612020	*PNPLA6*	AR	1st decade	Distal muscle atrophy, axonal motor neuropathy.
SPG40	Reserved	–	–	AD	–	Pure.
SPG41	11p14.1-p11.2	613364	Unknown	AD	Young adulthood (16–19)	Pure.
SPG42	3q25.31	612539	*SLC33A1*	AD	Variable (4–42)	Pure.
SPG43	19q12	615043	*C19orf12*	AR	1st decade	Distal muscle atrophy and weakness, neuropathy.
SPG44	1q42.13	613206	*GJC2*	AR	1st–2nd decade	Cerebellar ataxia, cognitive impairment.
SPG45/ SPG65	10q24.32-q24.33	613162	*NT5C2*	AR	2nd year	Mental retardation, delayed motor development.
SPG46	9p13.3	614409	*GBA2*	AR	Childhood (2–16)	Cerebellar ataxia, cataracts, mental retardation.
SPG47	1p13.2	614066	*AP4B1*	AR	Birth	Mental retardation, seizures, dystonia.
SPG48	7p22.1	613647	*AP5Z1*	AR	Adulthood (2–50)	Pure or complex with cognitive impairment or mental retardation..
SPG49	14q32.31	615031	*TECPR2*	AR	2nd year	Ataxia, brachycephaly, intellectual disability, hypoventilation, areflexia
SPG50	7q22.1	612936	*AP4M1*	AR	Birth	Mental retardation, seizures, cerebellar atrophy, strabismus.
SPG51	15q21.2	613744	*AP4E1*	AR	Birth	Mental retardation, seizures, cerebellar atrophy, hypotonia.
SPG52	14q12	614067	*AP4S1*	AR	Birth	Mental retardation, microcephaly, axial hypotonia, lack of speech development, stereotypic laughter, shyness.
SPG53	8p22	614898	*VPS37A*	AR	Infancy (1–2)	Kyphosis, cognitive impairment, delayed speech.
SPG54	8p11.23	615033	*DDHD2*	AR	Infancy (0–2)	Mental retardation, short stature, strabismus, telecanthus
SPG55	12q24.31	615035	*C12orf65*	AR	1st decade	Visual loss, optic atrophy, intellectual impairment, axonal neuropathy.
SPG56	4q25	615030	*CYP2U1*	AR	1st decade (0–8)	Pure or axonal neuropathy, dystonia, cognitive impairment.
SPG57	3q12.2	615658	*TFG*	AR	1st year	Optic atrophy, axonal and demyelinating sensorimotor neuropathy, muscle weakness and atrophy.
SPG58	17p13.2	611302	*KIF1C*	AR	Teenage	Pure or ataxia, dysarthria, distal muscle atrophy.
SPG59	15q21.2	603158	*USP8*	AR	Infancy	Nystagmus, mental retardation.
SPG60	3p22.2	612167	*WDR48*	AR	1st year	Nystagmus, neuropathy.
SPG61	16p12.3	615685	*ARL6IP1*	AR	Infancy	Acropathy, diffuse motor and sensory polyneuropathy.
SPG62	10q24.31	615681	*ERLIN1*	AR	Infancy	Pure.
SPG63	1p13.3	615686	*AMPD2*	AR	Infancy	Short stature, amyotrophy.
SPG64	10q24.1	615683	*ENTPD1*	AR	Infancy	Amyotrophy, cerebellar signs, intellectual disability.

Name	Locus	OMIM	Gene	Mode of inheritance	Average onset (range in years)	Distinguishing clinical features
SPG66	5q32	610009	*ARSI*	AR	Infancy	Amyotrophy, sensory motor and polyneuropathy.
SPG67	2q33.1	611655	*PGAP1*	AR	Infancy	Amyotrophy.
SPG68	11q13.1	604806	*FLRT1*	AR	Infancy	Optic atrophy, amyotrophy, neuropathy.
SPG69	1q41	–	*RAB3GAP2*	AR	Infancy	Dysarthria, cataract, deafness, intellectual disability.
SPG70	12q13.3	156560	*MARS*	AR	Infancy	Amyotrophy, contractures.
SPG71	5p13.3	615635	*ZFR*	AR	Infancy	Pure.
SPG72	5q31.2	615625	*REEP2*	AD/AR	Infancy	Pure.
SPG73	19q13.33	616282	*CPT1C*	AD	Adulthood (19–48)	Muscle atrophy, urinary dysfunction, delayed central sensory evoked potentials.
SPG74	1q42.13	616451	*IBA57*	AR	1st decade	Optic atrophy, axonal peripheral neuropathy, distal leg muscle atrophy.
SPG75	19q13.12	616680	*MAG*	AR	Early childhood	Optic atrophy, distal muscle atrophy, cognitive impairment, cerebellar signs, peripheral neuropathy.
SPG76	11q13.1	616907	*CAPN1*	AR	28 (19–39)	Cerebellar signs, sensory axonal neuropathy.
SPG77	6p25.1	611592	*FARS2*	AR	Before 5	Pure.
SPG78	1p36.13	617225	*ATP13A2*	AR	Juvenile or adulthood	Cognitive decline, strabismus, sensory-motor polyneuropathy.
SPG79	4p13	615491	*UCHL1*	AR	5–10	Blindness, cerebellar ataxia.
–	19p13.2	602378	*DNM2*	AD	Young Adulthood	Pure.
–	2p25.1	615759	*KIDINS220*	AD	1st year	Intellectual disability, nystagmus, obesity.
SPOAN	11q13.2	609541	*KLC2*	AR	Infancy	Optic atrophy and neuropathy.
–	6p21.1	–	*KLC4*	AR	Infancy	Hearing and vision loss, ataxic gait.
–	1q42.3	606897	*LYST*	AR	Adulthood (48–58)	Cerebellar ataxia, peripheral neuropathy.
–	5p15.2	256840	*CCT5*	AR	Infancy	Mutilating sensory neuropathy.
–	9p13.2	606489	*EXOSC3*	AR	Infancy	Cognitive disability, cerebellar signs, amyotrophy.
–	9q22.31	609797	*BICD2*	AD/AR	Infancy	Spinal muscular atrophy.
–	Mit	516060	*MT-ATP6*	AD	Adulthood (30–50)	Pain, axonal neuropathy.
–	1p36.22	607093	*MTHFR*	AR	Adulthood (29–50)	Polyneuropathy, behavioral changes, cognitive impairment, psychosis, seizures, leukoencephalopathy.
–	19p13.3	602662	*TUBB4A*	AD	Infancy	Pyramidal involvement, ataxia gait, dysdiadochokinesia.

SPG Spastic paraplegia, *AD* Autosomal dominant, *AR* Autosomal recessive

25.3.2 Prevalence

Prevalence of HSP varies widely among studies, probably due to a combination of factors, such as variable diagnostic criteria, epidemiological methodology, and population differences. Reported estimates vary from 0.1 to 9.6/100,000 in different series, 0.5–5.5/100,000 for dominant forms, and 0.0–5.3/100,000 for the recessive ones [9]. The most common dominant spastic paraplegia (SPG) in all series is SPG4, while SPG11 is the most frequent among the recessive HSPs.

25.3.3 Pathogenic Mechanisms

HSPs are among the most genetically heterogeneous diseases (Table 25.4). Many of the proteins involved act in the same cellular processes; nevertheless, the number of cellular mechanisms known to be affected continue growing and include: abnormal mitochondrial function, axonal transport dysfunction, alterations in lipid metabolism, abnormal DNA repair, alterations in membrane trafficking, organelle shaping and autophagy [12, 17]. Currently, no specific treatment exists to prevent, delay, or reverse progressive disability in patients with hereditary spastic paraplegia.

References

9. Coutinho P, Ruano L, Loureiro JL, Cruz VT, Barros J, Tuna A, Barbot C, Guimarães J, Alonso I, Silveira I, Sequeiros J, Marques Neves J, Serrano P, Silva MC (2013) Hereditary ataxia and spastic paraplegia in Portugal: a population-based prevalence study. JAMA Neurol 70:746– 755. https://doi.org/10.1001/jamaneurol.2013.1707

12. Lo Giudice T, Lombardi F, Santorelli FM, Kawarai T, Orlacchio A (2014) Hereditary spastic paraplegia: clinical-genetic characteristics and evolving molecular mechanisms. Exp Neurol 261:518–539. https://doi.org/10.1016/j.expneurol.2014.06.011

15. Hensiek A, Kirker S, Reid E (2014) Diagnosis, investigation and management of hereditary spastic paraplegias in the era of next-generation sequencing. J Neurol 262:1601–1612. https:// doi.org/10.1007/s00415-014-7598-y

17. Kara E, Tucci A, Manzoni C, Lynch DS, Elpidorou M, Bettencourt C, Chelban V, Manole A, Hamed SA, Haridy NA, Federoff M, Preza E, Hughes D, Pittman A, Jaunmuktane Z, Brandner S, Xiromerisiou G, Wiethoff S, Schottlaender L, Proukakis C, Morris H, Warner T, Bhatia KP, Korlipara LVP, Singleton AB, Hardy J, Wood NW, Lewis PA, Houlden H (2016) Genetic and phenotypic characterization of complex hereditary spastic paraplegia. Brain 139:1904–1918. https://doi.org/10.1093/brain/aww111

Acknowledgments

I want to thank deeply to my supervisors Dr. Isabel Alonso, Dr. Giovanni Stevanin and Prof. Dr. Jorge Sequeiros for having believed in me, accepting to guide me during these long years, providing this great learning opportunity and for the careful reading of this thesis.
To Isabel, for the work opportunity but mostly, for the large support that showed me from the beginning, the suggestions and solutions provided and for the kind help always there when needed. I am grateful for all these years working together. To Giovanni, for accepting to receive me, for all the pleasant help and guidance through this work, the suggestions and the teaching. To professor Jorge Sequeiros for accepting me in his laboratory a long time ago, leading me to continue my training and evolution.

I want to thank all the French team for the kind welcoming, allowing me to be part of such a nice group. More specifically, to Alexis Brice for receiving me in the laboratory, allowing me to have this learning opportunity. To Laure Raymond and Marie Coutelier for teaching me a lot, introducing me to the wonders of NGS and the lab with patience and sympathy. Marie, I loved your singing outbursts ;). To Mathilde Mairey and Liena Elsayed that always had kind words, were super helpful and friendly, proving to be there for me in times of need. Thank you a lot for that! To Federica Barreca, that although there for a short time, was also a fantastic companion. To Remi Valter that was really helpful and nice during my last visit. To Frédéric Darios and Alexandra Durr for the kind work suggestions. To Claire Pujol and Typhaine Esteves for your help and always nice smile. To Julien Branchu for the french lessons, although I showed to be a lousy student. To Christelle Tesson, Valérie Drouet, Maria Belen Lopez, Raphaël Matusiak and Maxime Jacoupy for accepting me and turning the lunch hour more interesting. To Lydia Guennec for all the bureaucratic help. To Sylvie Forlani and Christelle Dussert for enduring all my sample requests. To Justine Guégan for all the exome analysis, thank you for your explanations and patience to answer all my doubts. Finally, to Boris Keren for all the incredible help, without whom there would be no linkage analysis.

Agora em Português, quero muito agradecer a todos os membros da UnIGENe sem o qual não seria possível sobreviver neste trabalho todos estes anos, sem o seu apoio, ajuda e amizade. À Conceição Pereira, o meu pilar de sabedoria e manual de instruções para todas as minhas questões científicas e não científicas. Aprendi imenso contigo ao longo destes anos e não sei mesmo como o teria conseguido sem ti, muito obrigada pela imensa ajuda e paciência! À Diana Santos e Mariana Graça por toda a ajuda, mas também pelas palavras amigas com que posso sempre contar. À Marlene Quintas que embora distante, me faz muita

falta. Ao João Neto pelas piadas, boa disposição e vontade constante de ajudar. À Carolina Lemos e Miguel Ferreira pela sua alegria e força de vida. Às meninas Andreia Dias, Joana Silva, e agora Cátia Leal que embora tenham permanecido connosco pouco tempo, foram sempre uma agradável companhia.

Gostaria também de agradecer aos membros do CGPP, Patrícia Arinto, Ana Filipa Brandão, Ana Margarida Lopes, Rita Ferreira, Susana Sousa, Paulo Silva, Victor Mendes, Andreia Perdigão, Milena Paneque, Cristina Serrano e Susana Barbosa por toda a ajuda ao longo destes anos, desde placas de sequenciações a análises de exomas, a apoio logístico e informático. E claro, a muitos bolos partilhados durante todo este tempo :).

Gostaria de agradecer à equipa médica que permitiu que este trabalho pudesse ser realizado, principalmente à Doutora Paula Coutinho, grande força motriz do rastreio das famílias com paraparésias espásticas em Portugal. Ao Dr. José Leal Loureiro por todo o trabalho realizado durante estes anos e à sua colaboração com o nosso grupo. E também às Doutoras Eva Brandão e Paula Ribeiro pela sua disponibilidade e atenção e a todos os restantes médicos que colaboraram de alguma maneira neste trabalho.

Aos doentes e seus familiares pela sua disponibilidade e consentimento para estudo do material biológico, sem o qual a realização deste trabalho não seria possível.

Gostaria também de agradecer às agências financiadoras sem as quais não seria possível concretizar este projeto e, principalmente à FCT, que com a atribuição da minha bolsa de doutoramento, reconheceu o potencial deste trabalho.

Por fim, mas sempre em primeiro, queria agradecer à minha família. Ao Ricardo, a minha nova família e fonte de força e inspiração que tem estado sempre comigo a apoiar-me nos bons e maus momentos, a encorajar-me quando começo a perder o foco e pela enorme paciência que tem tido comigo desde sempre. Muito obrigado amor, e que uma nova etapa comece ;)! Aos meus pais que sempre me apoiaram e incentivaram a evoluir e lutar pelos meus sonhos, e me ajudaram quando foi preciso ultrapassar os momentos mais difíceis. Principalmente à minha mãe por me incentivar a acabar com as suas constantes perguntas de “e quando acaba isso?” ;). Ao meu irmão e à minha avó por todo o carinho, força e paciência que têm comigo. Queria também agradecer à minha nova família por casamento que me recebeu de braços abertos e à minha família adotada, Evo, por todo o apoio que me deram ao longo destes anos. Muito obrigada a todos!

The work presented in this thesis was performed at the Instituto de Biologia Molecular e Celular, Instituto de Investigação e Inovação em Saúde (IBMC - I3S, University of Porto), Porto, Portugal and at the Institute of Institut du Cerveau et de la Moelle épinière, Paris, France. This study was funded in part by the FCT (Fundação para a Ciência e Tecnologia, grant FCT-ANR/BEX-GMG/0008/2013, to IA) and by the French National Agency for Research (SPATAX-QUEST, to GS), the European Union (Omics call of the 7th framework program, NEUROMICS, to AB; the E-rare program, NEUROLIPID and PREPARE, to GS), the VERUM Foundation (to AB and GS) and benefited from the program “Investissements d’Avenir” ANR-10-IAIHU-06 (to the ICM). SM is the recipient of a fellowship (SFRH/BD/87189/2012) from the FCT supported by POPH/FSE funding.