Research Programs Unit – Stem Cells and Metabolism Research Program
and
Doctoral Programme in Biomedicine
Faculty of Medicine
University of Helsinki
Finland

GENETICS AND MECHANISMS OF RARE INHERITED NEUROMUSCULAR DISORDERS – A FOCUS ON NEUROFILAMENT LIGHT

Markus Sainio

ACADEMIC DISSERTATION

To be presented, with the permission of the Faculty of Medicine of
the University of Helsinki, for public examination in lecture hall Athena 107,
Siltavuorenpenger 3, on 6 September 2022, at 13.00.

Helsinki, Finland 2022

Supervised by
Professor Henna Tyynismaa, PhD
STEMM, Research Programs Unit, Faculty of Medicine
University of Helsinki

Docent Emil Ylikallio, MD, PhD
STEMM, Research Programs Unit, Faculty of Medicine
University of Helsinki

Reviewed by
Associate Professor Mika Martikainen, MD, PhD
Department of Clinical Medicine, Clinical Neurosciences, Faculty of Medicine
University of Turku

Docent Šárka Lehtonen, PhD
A.I.Virtanen Institute for Molecular Sciences
University of Eastern Finland

Opponent
Research director Pascale Bomont, PhD
INSERM
University of Lyon

Custos
Professor Henna Tyynismaa, PhD
STEMM, Research Programs Unit, Faculty of Medicine
University of Helsinki

ISSN 2342-3161 (print)
ISSN 2342-317X (online)
ISBN 978-951-51-8367-5 (print)
ISBN 978-951-51-8368-2 (online)
Dissertationes Scholae Doctoralis Ad Sanitatem Investigandam Universitatis Helsinkiensis

The Faculty of Medicine uses the Ouriginal system (plagiarism recognition) to examine all doctoral dissertations.

Unigrafia
Helsinki 2022

TIIVISTELMÄ

Perinnöllisten neuromuskulaaritautien ryhmä on kokoelma lihas- ja hermoperäisiä sairauksia, jotka aiheuttavat liike oireita. Potilaiden kliinisten oireiden perusteella voi olla vaikea erottaa johtuvatko liike oireet hermojen vai lihasten toimintahäiriöstä, sillä molemmat johtavat lihasheikkouteen. Hoitojen kehittämiseksi oireiden perimmäiset syyt ja molekyylitason mekanismit tulee selvittää. Nykyään perinnöllisen sairauden tautimekanismin selvityksessä voidaan käyttää geneettisiä uuden sukupolven sekvensointimenetelmiä, kuten eksomisekvensointia. Näillä menetelmillä pystytään jopa sekvensoimaan potilaan koko genomi ja selvittämään missä geenissä oleva virhe eli mutaatio johtaa taudinkuvaan. Taudin oireiden geneettisen syyn selvitys ei kuitenkaan riitä hoitojen kehittämiseen. Geneettisen muutoksen aikaansaava molekyylitason tautimekanismi tulee selvittää juuri niissä solutyypeissä, jotka kyseisessä taudissa rapistuvat. Tautimekanismin yksityiskohtainen selvitys voi paljastaa prosesseja, joihin voisi vaikuttaa lääkinnällisesti, ja näin hidastaa taudin etenemistä tai estää sen pahentumista.

Tämän väitöskirjan tavoitteina on selvittää harvinaisten perinnöllisten neuromuskulaaritautien genetiikkaa, arvioida uuden sukupolven sekvensointimenetelmien tehokkuutta tässä tautiryhmässä ja mallintaa *NEFL* (*neurofilament light*) geenin nonsense-mutaatiosta johtuvan aksonirappeumansairauden, Charcot-Marie-Tooth neuropatian (CMT), tautimekanismeja kantasoluista erilaistetuilla liikehermosoluilla.

Harvinaisten perinnöllisten neuromuskulaaritautien sekvensointiprojektissa 27 % potilaista sai taudilleen geneettinen syyn eksomisekvensoinnilla. Potilasryhmä koostui 100 potilaasta, joilla oli vaihtelevia taudinkuvia, kuten myopatiaa, neuropatiaa, spastisuutta, ataksiaa ja parkinsonismia. Löysimme kaksitoista uutta mutaatiota kolmeltatoista potilaalta geeneistä, jotka on aiemmin liitetty neuromuskulaaritauteihin. Tutkimukseni osoittaa, että eksomisekvensointi on tehokas menetelmä neurmuskulaaritautien geneettisten syiden selvittämiseksi aikuispotilailla, joilla epäillään perinnöllistä tautia. Tämän lisäksi kuvailin uuden oireenkuvan suomalaisilla myopatiapotilailla, joilla on mutaatioita *PYROXD1 (Pyridine Nucleotide-Disulphide Oxidoreductase Domain 1)* geenissä. Heidän oireensa alkoivat aikuisiässä ja taudinkuva vastaa limb-girdle tyyppistä lihasdystrofiaa (LGMD), aiemmin kuvatuilla potilailla lihasoireet alkoivat varhaisessa lapsuudessa eivätkä muistuttaneet LGMDtä.

NEFL geenissä esiintyvät nonsense-mutaatiot johtavat peittyvästi periytyvään aksonirappeumasairauteen, jossa oireet alkavat tavallisesti lapsuudessa. Peittyvästi periytyvien *NEFL* muutosten tautimekanismi on yhä selvittämättä. Vallitsevasti periytyvien *NEFL* muutosten ajatellaan aiheuttavan aksonirappeumaa proteiinin epätavallisen kasautumisen seurauksena. Löysin uuden peittyvästi periytyvä muutoksen p.R367X *NEFL* geenissä, joka aiheuttaa varhain alkavaa aksonirappeumaa. Tutkin uutta muutosta potilaan soluista tehdyillä kantasoluilla, jotka erilaistin liikehermosoluiksi. Osoitin että vallitsevasti periytyvien muutosten vastaisesti, tämä uusi peittyvästi periytyvä muutos johtaa *NEFL* lähetti-RNA molekyylin hajotuksen kautta proteiinin (NFL) puutokseen. Mallintaakseni NFL proteiinin puutosta tehokkaasti, poistin *NEFL* geenin terveistä kantasoluista

geenieditoinnilla. Osoitin geenieditoiduilla ja *NEFL* potilassoluilla, että NFL proteiinin puutos johtaa CMT neuropatiaan, vaikkakin NFL proteiinin puutos ei häirinnyt kantasoluista erilaistettujen liikehermosolujen kypsymistä tai aksonien pituuskasvua. Tämän lisäksi solut pystyivät tuottamaan neurofilamentteja ilman NFL proteiinia. Näiden filamenttien rakenne oli kuitenkin muuttunut ja niitä esiintyi harvemmin. Tutkimukseni paljasti NFL proteiinin puutoksen johtavan kapeampiin aksoneihin, alentuneisiin elektrofysiologisiin vasteisiin synapsissa ja mitokondrioiden liikkeen lisääntymiseen erilaistetuissa liikehermosoluissa.

Aiemmat kliiniset tutkimukset *NEFL* neuropatiapotilailla ja kokeet *Nefl* poistogeenisillä hiirillä ovat viitanneet NFL proteiinin puutoksen johtavan yllä mainittuihin fenotyyppeihin. Aksonit, ja etenkin myelinoidut aksonit, ovat kapeampia potilailla, joilla on peittyvästi periytyvä neuropatia *NEFL* geenin mutaation seurauksena. Kapeat aksonit johtavat aktiopotentiaaleja hitaammin hermosolun soomasta aksonien kärkeen, joka johtaa heikompaan lihasvasteeseen ja oireisiin. Osoitin ensimmäistä kertaa ihmissoluilla, että NFL puutos johtaa alentuneisiin synaptisiin elekrofysiologisiin vasteisiin, tämä vaikutus on aiemmin todettu vain *Nefl* poistogeenisillä hiirimalleilla. Näiden hiirimallien avulla on osoitettu NFL proteiinin tukevan glutamaattireseptorien toimintaa synapsissa. Sillä molemmissa malleissa elektrofysiologiset vasteet ovat alentuneet, myös ihmisillä NFL saattaa tukea synapsin toimintaa. NFL toimintaa synapseissa tulisi tutkia tarkemmin, jotta sen rooli aksonirappeumassa pystyttäisiin todentamaan. *Nefl* poistogeenisten hiiritutkimusten ja tämän väitöskirjan NFL puutosmallien tulosten perusteella voimme todeta NFL proteiinin toimivan mitokondrioiden ankkurina aksoneissa. Molemmissa malleissa NFL puutos johtaa mitokondrioiden lisääntyneeseen liikkeeseen. Tämä lisääntynyt liike voi vähentää paikallaan olevien aktiivisten mitokondrioiden määrää kriittisissä solunsisäisissä sijainneissa, johtaen paikallisiin energiavajeisiin ja aksonin degeneraatioon.

Tämä työ osoittaa, että kantasolupohjaisia hermosolumalleja pystytään hyödyntämään aksonirappeumatautien mallinnuksessa. Tässä väitöskirjassa esitettyjä patologisia molekyylitason mekanismeja perinnöllisissä aksonirappeumasairauksissa tulisi jatkossa tutkia lisää, jotta voisimme paikantaa lääkinnällisiä kohteita taudin oireiden lievittämiseen tai etenemisen pysäyttämiseksi.

ABSTRACT

Hereditary neuromuscular disorders are a heterogenous group of diseases affecting motor nerves and musculature. The clinical distinction between nerve or muscle originating disorders can be difficult since both neuronal loss and muscle disruption lead to muscle weakness and atrophy. The underlying genetic cause of a neuromuscular disease and its pathological molecular mechanism need to be unraveled to develop disease modifying treatments. Fortunately, next-generation sequencing (NGS) technologies have revolutionized the molecular diagnostics of hereditary diseases. High throughput technologies can be used to efficiently sequence all possible pathogenic variants in an individual's genome to discover the affected gene. However, the gene variant behind the disorder does not directly indicate the mechanism of disease. To discover treatable processes, the cellular and molecular alterations caused by the mutant gene must be modelled in appropriate systems.

In this dissertation I aimed to discover genetic causes behind rare hereditary neuromuscular disease in adult patients, evaluate the efficacy of NGS in this patient group, and model *NEFL* (*neurofilament light*) nonsense variants that cause Charcot-Marie-Tooth (CMT) disease with induced pluripotent stem cell (iPSC) derived motor neurons (iPSC-MN) to discover treatable mechanisms of disease.

In a cohort of 100 adult patients, we discovered pathogenic variants in 27% of patients with varied hereditary disorders such as myopathy, neuropathy, spasticity, ataxia, and parkinsonism using clinical whole-exome sequencing (WES). Twelve of the variants in 13 patients were novel changes in known disease-associated genes, thus we expanded the genetic causes and phenotypes of these disorders. I showed WES to be efficient and potentially cost-effective in the molecular diagnosis of adult patients with suspected hereditary origin of disease. In addition, we discovered Finnish patients with previously uncharacterized phenotypes linked to *PYROXD1* variants. The Finnish patients had a late-onset limb-girdle muscular dystrophy (LGMD)-like phenotype in contrast to the previously described congenital myopathy with myofibrillar changes.

Nonsense variants in the *NEFL* gene cause a severe, early-onset form of axonal CMT through an uncharacterized mechanism. Studies on dominant *NEFL* variants indicate the axon degeneration to be caused by mutant protein aggregation, however, the disease mechanism behind nonsense cases is still elusive. I characterized a novel pathogenic recessive nonsense variant p.R367X in *NEFL* causing CMT. I showed the novel variant to lead to nonsense-mediated decay of mRNA and to neurofilament light polypeptide (NFL) protein loss in iPSC-MN of the patient. We generated isogenic *NEFL* knockout cell lines from control iPSC to more accurately assay the effects of NFL loss. I characterized the patient and isogenic gene-edited iPSC-MN devoid of NFL and described specific molecular alterations in the neurons. I showed the absence of the integral neurofilament (NF) protein NFL to be causative in recessive *NEFL* CMT but not to affect iPSC-MN differentiation or the growth of elaborate axonal connections in patient or isogenic gene-edited models. Furthermore, the neurons were able to produce NFs without compensation from other

intermediate filament proteins. However, there were fewer NFs and their composition was altered by NFL absence and reduction of neurofilament heavy polypeptide (NFH). The loss of NFL reduced axon diameter, decreased amplitude of miniature post-synaptic currents (EPSCs) and increased movement of axonal mitochondria in iPSC-MNs.

Clinical studies on patients and *Nefl* knockout mouse models have shown similar mechanisms of disease as our iPSC-MN model. The caliber of large axons is also reduced in patients with *NEFL* nonsense mutations. In the patients the reduction of axon size and loss of large myelinated axons leads to decreased nerve conduction velocity. The synaptic function of NFL has not been previously described in human models, but NFL has been shown to bind glutamatergic receptors in mice. The reduced amplitude in EPSCs suggests a synaptic regulatory role of NFL also in humans warranting further research. The loss of NFL in the synapse could reduce the excitability of spinal motor neurons leading to reduced activity. Increase of mitochondrial movement in axons without NFL adds to previous research on mice indicating NFs and NFL working as scaffolds for organelles in axons. The increase in mitochondrial movement could reduce the amount of correctly localized active mitochondria and lead to local metabolic deficiencies predisposing axons to degeneration.

My iPSC-MN modelling studies show the usability of stem cell-based models in studying the mechanisms of axonal neuropathy and its pathological processes. Further research is needed to pinpoint the translatable pathological alterations in the human axons that could be targeted for treatments in patients.

CONTENTS

LIST OF ORIGINAL PUBLICATIONS

I Sainio M.T, Aaltio J, Hyttinen V, Kortelainen M, Ojanen S, Paetau A, Tienari P, Ylikallio E, Auranen M, Tyynismaa H. Effectiveness of clinical exome sequencing in difficult-to-diagnose neurological diseases of adults. Acta Neurologica Scandinavica 2021;145(1), 63–72.

II Sainio M.T, Välipakka S, Rinaldi B, Lapatto H, Paetau A, Ojanen S, Brilhante V, Jokela M, Huovinen S, Auranen M, Palmio J, Friant S, Ylikallio E, Udd B, Tyynismaa H. Recessive PYROXD1 mutations cause adult-onset limb-girdle-type muscular dystrophy. Journal of Neurology 2019;266(2), 353-360.

III Sainio M.T, Ylikallio E, Mäenpää L, Lahtela J.S.P, Mattila P.M, Auranen M, Palmio J, Tyynismaa H. Absence of NEFL in patient-specific neurons in early-onset Charcot-Marie-Tooth neuropathy. Neurology Genetics 2018;4(3), e244.

IV Sainio M.T, Rasila T, Molchanova S, Järvilehto J, Torregrosa-Muñumer R, Harjuhaahto S, Pennonen J, Huber N, Herukka S, Haapasalo A, Zetterberg H, Taira T, Palmio J, Ylikallio E, Tyynismaa H. Neurofilament light regulates axon caliber, synaptic activity and organelle trafficking in cultured human motor neurons. Frontiers in Cell and Developmental Biology 2022; 9:820105.

In addition, unpublished results are presented.

ABBREVIATIONS

ABCD1	*ATP Binding Cassette Subfamily D Member 1*
ACMG	American College of Medical Genetics and Genomics
AMPA	Alpha-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid
AMX	Amlexanox
ANO5	*Anoctamin 5*
ANOVA	Analysis of variance
AP-3	Adaptor protein-3
Ascl1	*Achaete-Scute Family BHLH Transcription Factor 1*
ATP	Adenosine triphosphate
BAG3	*BAG Cochaperone 3*
BDNF	Brain derived neurotrophic factor
BMP	Bone morphogenic protein
Bp	Base pair
Brn2	*POU Class 3 Homeobox 2*
Cas9	CRISPR associated protein 9
CES	Clinical exome sequencing
Chat	Acetylcholine transferase
CHCHD10	*Coiled-Coil-Helix-Coiled-Coil-Helix Domain Containing 10*
CHX	Cycloheximide
CMAP	Compound muscle axon potential
CMT	Charcot-Marie-Tooth disease
c-Myc	*MYC Proto-Oncogene, BHLH Transcription Factor*
CNTF	Ciliary neurotrophic factor
CRISPR	Clustered regularly interspaced short palindromic repeats
crRNA	CRISPR RNA
DAPT	γ-secretase inhibitor
DBS	Double strand break
DBX	*Developing brain homeobox*
DCAF8	*DDB1 And CUL4 Associated Factor 8*
ddNTP	Dideoxynucleotides
DG	Dystrophin glycoprotein complex
DNA	Deoxyribonucleic acid
DYSF	*Dysferlin*
ECM	Extracellular matrix
EMG	Electromyography
ENMG	Electro Neuro-Myography
ER	Endoplasmic reticulum
ES	Embryoid stem cell
FACS	Fluorescence activated cell sorting
FAD/H	Flavin adenine dinucleotide

FDH	Finnish disease heritage
FGF	Fibroblast growth factor
FIMM	Finnish Institute of Molecular Medicine
GAN	*Giant axonal neuropathy*
GDF11	*Growth differentiation factor 11*
GDNF	Glial derived neurotrophic factor
GFP	Green fluorescent protein
GJB1	*Gap Junction Protein Beta 1*
GLI	*GLI-Kruppel family protein*
GRIN1	*Glutamate Ionotropic Receptor NMDA Type Subunit 1*
GWAS	Genome-wide association studies
HDR	Homology-directed repair
HeLa	Henrietta Lacs derived cancer cell line
HMN	Hereditary motor neuropathy
HSN	Hereditary sensory neuropathy
HSP	Hereditary spastic paraplegia
HSPB1	Heat Shock Protein Family B (Small) Member 1
HSPB8	Heat Shock Protein Family B (Small) Member 8
IF	Intermediate filament
INA	α-internexin
iPSC	Induced pluripotent stem cells
iPSC-MN	Induced pluripotent stem cell derived motor neuron
IRX3	*Iroquois related homeobox 3*
ISL1	*ISL1 transcription LIM homeodomain*
Klf4	*Kruppel Like Factor 4*
KO	Knock-out
LGMD	Limb girdle muscular dystrophy
LHX3	*LIM homeobox 3*
LMC	Lateral motor column
Lmx1a	*LIM Homeobox Transcription Factor 1 Alpha*
mEPSC	Miniature excitatory post-synaptic current
MFN2	*mitofusin 2*
MIRAS	Mitochondrial recessive ataxia syndrome
MN	Motor neuron
MNX1/HB9	*Motor neuron and pancreas homeobox 1*
MPZ	*Myelin protein zero*
MRI	Magnetic resonance imaging
mRNA	Messenger ribonucleic acid
MT-ATP6	*Mitochondrially Encoded ATP Synthase Membrane Subunit 6*
Myt1l	*Myelin Transcription Factor 1 Like*
NADH	Nicotinamide adenine dinucleotide
NCV	Nerve conduction velocity
NEFL	*Neurofilament light*
NEUROG2	*Neurogenin 2*

NF	Neurofilament
NFH	Neurofilament heavy polypeptide
NFL	Neurofilament light polypeptide
NFM	Neurofilament medium polypeptide
NGS	Next-generation sequencing
NHEJ	Non-homologous end joining
NKX2	*NK2 homeobox*
NKX6	*NK6 homeobox*
NMD	Nonsense mediated decay
NPHS1	*Nephrin*
Nurr1	*Nuclear Receptor Subfamily 4 Group A Member 2*
NVC	Nerve conduction velocity
Oct3/4	*POU Class 5 Homeobox 1*
ODN	Oligonucleotide
OLIG2	*Oligodendrocyte transcription factor 2*
OMIM	Online Mendelian inheritance in man
PAM	Protospacer-adjacent motif
PAX	*paired box*
PCR	Polymerase chain reaction
PDC	Pyruvate dehydrogenase complex
PDK3	Pyruvate dehydrogenase kinase 3
PFA	Paraformaldehyde
pMN	Motor neuron progenitor
pmn mice	Tubulin binding factor E mouse model
PMP22	*Peripheral Myelin Protein 22*
PNDR1	Class 1 nuclear-cytoplasmic pyridine nucleotide-disulphide reductase
POLG	*Polymerase gamma*
PRPH	*Peripherin*
PTC-124	Ataluren
PYROXD1	*Pyridine nucleotide-disulphide oxidoreductase domain 1*
RA	Retinoic acid
ROCKi	RHO/ROCK pathway inhibitor
SAG	Smoothened agonist
Shh	Sonic Hedgehog Signaling Molecule
SMAD	Family of proteins binding TGF-B receptors
SNP	Single nucleotide polymorphism
SOX2	*SRY-box transcription factor 2*
SOX3	*SRY-box transcription factor 3*
TALEN	Transcription activator-like effector nuclease
TDP-43	*TAR DNA Binding Protein*
TGF-β	Transforming growth factor beta
TMD	Tibial muscular dystrophy
TRIM2	*Tripartite Motif Containing 2*

TTN	Titin
UFL	Unit-length filaments
VUS	Variants of unknown significance
WES	Whole exome sequencing
WGS	Whole genome sequencing
Wnt	Wingless
ZFN	Zinc finger nuclease

1 INTRODUCTION

An estimate of 30 million Europeans will suffer from a rare disease in the course of their life (Fernandez-Marmiesse et al., 2017). A rare disease affects fewer than 50 in 100 000 (Boycott et al., 2013). Adult-onset hereditary neuromuscular diseases, which include disorders such as Charcot-Marie-Tooth disease (CMT) and Limb-girdle muscular dystrophy (LGMD), are a large clinically and genetically heterogenous group of rare diseases. CMT is often referred as the most common hereditary neuromuscular disorder with an estimated prevalence of 40 in 100 000 (Skre, 1974). A similar prevalence, 37 per 100 000, is estimated for hereditary muscle diseases in England where LGMD is seen approximately in 3 per 100 000 (Norwood et al., 2009). Hereditary neuromuscular diseases cause symptoms by affecting different tissues in the central and peripheral nervous system. In myopathies, such as LGMD (Straub et al., 2018) and neuropathies such as CMT (Pipis et al., 2019), peripheral tissues are affected. Central disorders, which have overlapping symptoms and even genetics with peripheral neuromuscular disease, include ataxia (Parodi et al., 2018) with an estimated prevalence of 6 per 100 000 (Hellberg et al., 2019), and hereditary spastic paraplegia (HSP) (Shribman et al., 2019) seen in 2-3 per 100 000 (Hellberg et al., 2019). To develop disease altering treatments for this diverse group of disorders, the clinical features, genetic etiology, cellular, and molecular mechanisms of the disease must be uncovered and described.

Hereditary neuromuscular disorders cause progressive muscle weakness and atrophy, as well as difficulties in fine motor control (Pipis et al., 2019). As the diseases progress additional symptoms arise, such as pain, numbness, stiffness etc. (Klein, 2020; Straub et al., 2018). Muscular and neuronal disorders have symptoms that are occasionally hard to discern from one another. Muscle weakness can originate from neuronal signal transduction difficulties, muscle contraction problems or muscle dystrophy. For example, neuronal denervation of muscle can be the cause of muscle atrophy, and neuronal degeneration can cause denervation and muscle atrophy. In adult and elderly cases with a suspected genetic origin of disease, symptoms can be further masked by the patient's acquired phenotypic features. Additionally, the genetic burden of risk factors can bring additional complexity to patient diagnosis and genetics (Bis-Brewer et al., 2020; Posey et al., 2017).

In clinical diagnostics, genetic analysis through next-generation sequencing (NGS) technologies have revolutionized the identification of causes of hereditary neuromuscular diseases (Adams & Eng, 2018; Córdoba et al., 2018; Splinter et al., 2018). Patient diagnosis is no more only reliant on clinical interpretation of symptoms and sequencing of single genes (Vissers et al., 2017). Now a whole exome (WES) or genome (WGS) of an individual can be sequenced to discover variants that are possibly responsible for the disease in a high throughput manner (Gorcenco et al., 2020). In WES all exons of genes are sequenced, which allows the discovery of variants in coding areas of the genome. In WGS exons, introns, and, as the name implies, the rest of the genome is sequenced. Currently, WES is a more feasible

sequencing technology in the clinic since the amount of data is more approachable, current databases contain large amounts of previously validated pathogenic variants in exons and the frequency of variants in the exons is available from large cohorts from multiple populations (Schwarze et al., 2018). In the case of WGS, the amount of data can be difficult to handle and databases with non-coding pathogenic variants are lacking (Adams & Eng, 2018). The early use of WES can help clinicians to pinpoint the genetic and molecular diagnosis for patients with hereditary neuromuscular disorders (Stark et al., 2019). In my thesis, I assess the feasibility of clinical whole exome sequencing (CES) in a cohort of adult-onset neuromuscular patients. I show CES to be efficient in finding pathogenic and likely pathogenic variants in a diverse cohort of patients.

To develop treatments for the neuromuscular disorders, the molecular and cellular mechanism of disease should be investigated in a relevant model system. Researchers must uncover the chain of events in the cell that causes a disease with a given mutation in order to intervene with the pathogenic process. Animal models are usable in disease modelling but lack in translatability from model to human disorders. The animals might display different phenotypes compared to humans (Anderson et al., 2019) or no pathological phenotype at all (Zhu et al., 1997). For example, the majority of preclinical cancer trials conducted on animal models with positive results are not replicated in later clinical trials on humans (Perlman, 2016; Richmond & Yingjun, 2008). In modelling tissue specific diseases, one must study the tissues affected to accurately mimic the disease state present in the patient's respective tissue. For example, in neuropathy modelling human *in vitro* models of non-neural cells are lacking in accuracy. Fortunately, somatic cells, such as fibroblasts, can be reprogrammed into induced pluripotent stem cells (iPSC) (Takahashi et al., 2007; Takahashi & Yamanaka, 2006) that can then be differentiated into motor neurons (Maury et al., 2015) to investigate neuron specific disease mechanisms, or any other cell type to relevant for the disorder. Through iPSC we can model mutations causing motor neuron axon degeneration in the affected cell type on a dish.

Since iPSC proliferate indefinitely, there is theoretically no limit in how much neuronal material can be differentiated. Hence, researchers can study the molecular mechanism of disease with numerous methods. Messenger-RNA level alterations can be studied with transcriptomics (Sainio et al., 2018), protein level composition of neurons with proteomics or specific proteins with immunoblotting (Juneja et al., 2018), axonal transport of organelles with live cell imaging (Van Lent et al., 2021) etc. Upon discovery of a druggable pathway, researchers can investigate such drugs with ease in a human cell type specific model (Perez-Siles et al., 2020). It is still early to compare the translatability of these novel iPSC-derived models to animal models or to human patients, but current research is promising.

The research on rare hereditary disorders has been often limited to rare patient specific cells carrying the causative mutations. Accurate genetic editing to produce isogenic cell lines has been expensive and time consuming (Torres-Ruiz & Rodriguez-Perales, 2017) before the

development of clustered regularly interspaced short palindromic repeats (CRISPR)-Cas9 technology. With CRISPR-Cas9 control iPSC lines can be edited to carry a pathogenic mutation, or the disease mutations can be corrected in a patient iPSC cell line with relative low cost and time (Anzalone et al., 2019). Therefore, currently researchers can model hereditary diseases in the affected cell type comparing to isogenic controls with only the mutation under study differing between them.

In my thesis I will discuss the discovery of pathogenic variants in patients with rare hereditary neuromuscular disorders and the modelling of neuropathy caused by *neurofilament light* nonsense mutations in iPSC-derived motor neurons.

2 REVIEW OF THE LITERATURE

2.1 Human genome, genetics and inheritance of variants

2.1.1 From DNA to translation

Humans have a diploid chromosome with 22 homologous chromosome pairs (autosomes) and the sex chromosomes XY (male) or XX (female) equaling 44 + XY/XX. One allele of each chromosome pair is inherited from the mother and one from the father. In addition, humans have a separate small mitochondrial genome (16,569bp) inherited from the mother. Chromosomes consist of tightly packed deoxyribonucleic acid (DNA), which is coiled around histone proteins. DNA is formed from two antiparallel strings of nucleotides (bases); the pyrimidines cytosine (C) and thymine (T); and purines adenine (A) and guanine (G). The molecules contain a circular core pentose sugar. The first carbon in the circular sugar molecule is bound to one of the four bases, the third carbon (3') has a hydroxyl group and the fifth (5') contains a phosphate group. Bases in two antiparallel strands connect with each other through Watson-Crick base pairing; adenosine (A) binds thymine (T), and guanine (G) binds cytosine (C), and the sugar moieties of the nucleotides create a backbone, which twists creating a double helix (Watson & Crick, 1953). The human genome contains approximately 3 billion nucleotide pairs (3.1Gb), currently estimated to code for 21 000 genes and proteins (Lander et al., 2001).

The human genome is divided into euchromatin containing genes and gene regulatory areas, and the heterochromatin containing silent and repetitive genomic areas. Protein coding genes are made of exons and introns. The exons contain the genetic code for the protein, and introns have regulatory information (Morrison & Thakur, 2021). When a gene is transcribed, an RNA polymerase binds the gene promoter and copies the genetic information into a single stranded messenger ribonucleic acid (mRNA) molecule. DNA transcription, as well as replication, is always performed from 5' to 3' direction. The DNA antisense strand is used as the template to create mRNA with the same sequence as the sense strand. The mRNA is formed of the same nucleotides as DNA A, C, G, but thymine is replaced by uracil (U). The mRNA molecule is processed in the nucleus; introns are removed by splicing, a poly-A tail is added to the 3'end and the 5'end is capped with a modified guanine moiety (Djebali et al., 2012). In the cytoplasm the processed mRNA is translated into a protein by the ribosome from amino acids loaded onto transfer RNA (tRNA) molecules. Nucleotides in the mRNA form three nucleotide long sequences called codons. The 20 amino acids that form all proteins are loaded into tRNA complexes with three nucleotides forming anticodons, thus the DNA and in extension mRNA code is translated into an amino acid code. Different codons code for different amino acids. When the codon in the mRNA meets the anticodon in the tRNA, a specific amino acid is added to the growing polypeptide chain. When the ribosome encounters a stop-codon in the mRNA, the translation is terminated.

In case of a premature stop codon, caused by a nonsense mutation, the ribosome and nonsense-mediated decay (NMD) machinery recognize the early stop codon by increased distance from the poly-A tail. The recognition of the premature stop codon will initiate NMD, which leads to the degradation of the mutant mRNA (Nagel-Wolfrum et al., 2016).

2.1.2 Variants in DNA

The human genome is subject to endogenous and exogenous forces that can cause damage to the DNA. Exogenous sources of DNA damage, the rarer of the two, include ionizing radiation and mutagenic chemicals. In most cases alterations in the genome are caused by endogenous factors, DNA replication errors or intracellular chemical damage. Replication errors are the result of enzymatic imperfection since no enzyme is perfect. The human DNA polymerase has 3'-5' proofreading capacity to fix errors in replication, but again no enzyme works with 100% efficiency, resulting in mistakes during the numerous rounds of DNA replication in one's life. Endogenous chemical damage can be caused by hydrolytic and oxidative sources, as well as abnormal DNA methylation. DNA damage or errors not repaired become mutations or variants (Roos et al., 2016).

All humans have approximately 3 million variants in their genome when compared to a compiled reference genome. The genetic variants can alter the amount of DNA or change the code without affecting DNA content. The variants can be large chromosomal alterations, such as trisomy of chromosome 21 causing Down syndrome, or small single nucleotide variants, changing only one nucleotide, or anything in between. Table 1 describes the types of genetic variation. The majority of variants are benign changes in single nucleotides or small insertions or deletions. The different variations in the human genome can be common in the population or unique to the individual. Common single nucleotide variants present in over 1% in the population have been termed as single nucleotide polymorphisms (SNPs). Variants are categorized as heterozygous, present in one allele, and homozygous, present in both alleles. When a different variant is present in two locations in the same gene in different alleles, the genotype is referred as compound heterozygous. Additionally, males with a variant in their only X-chromosome are termed hemizygous (Altshuler et al., 2010).

Table 1. Types of genetic variation.

DNA variant	Symbol	Example	Definition
Substitution	**>**	**ATCG->ATCC**	**One nucleotide is changed to another**
Insertion	**Ins**	**ATCG->ATCGA**	**Addition of a sequence**
Deletion	**Del**	**ATCG->ATC-**	**Removal of a sequence**
Inversion	**Inv**	**ATCG->CGAT**	**Original sequence is changed to the reverse-complement of itself**
Duplication	**Dup**	**ATCG->ATCGCG**	**Insertion of an existing sequence directly 3' of the original**
Conversion	**Con**	**ATCG->AAGTG**	**Deletion and specific insertion of a sequence from another genetic loci**
Deletion/insertion	**Del/ins**	**ATCG->AGTAG**	**Deletion of a sequence and addition of a non-complimentary sequence**

Changed nucleotides are highlighted.

Genetic variation is inherited from one's parents, one allele from each, and their variants are inherited from their parents. Historically, humans did not mix globally but concentrated into geographical regions. The people in the separated regions reproduced within the region. This created genetically separated populations around the globe with specific sets of genetic variants through the founder effect. In human history, populations have also been subject to genetic bottlenecks in which an environmental influence has reduced the local population dramatically causing all future generations to be closely related (Bergström et al., 2021; Nielsen et al., 2017). The Finnish population has gone through isolations and founder effects causing a unique Finnish disease heritage (FDH) (Norio, 2003a).

Genes are inherited in chromosomes. Chromosomes rearrange during meiosis where segments of the chromosome pairs are interchanged with each other. Therefore, variants in the genome are not inherited independently but as part of chromosomes or sections in chromosomes, termed as haplotypes. People in the same population commonly share many of these haplotypes with each other. The haplotypes have variants that can be used as markers of the specific haplotype. Variants inside the haplotype are in a linkage disequilibrium since the variants are seen together inside a haplotype more commonly than would be expected by statistics if all variants would be separately inherited (Henn et al., 2015).

2.1.3 Consequences of DNA variants

Since approximately 1% of the genome codes for protein, most variants reside in non-coding areas. The variants in exons, on an amino acid level, are categorized with their effect on the protein. Missense mutations change the amino acid, synonymous variants are changes in the coding region that do not change the amino acid, nonsense mutations change the codon to a stop-codon, and frameshift variants disrupt the reading frame, which changes the whole amino acid code from the variant forward likely causing a stop-codon to appear either prematurely or upstream in the gene. Changes in exon-intron boundaries can affect mRNA

splicing and are therefore termed as splice-site variants. Table 2 summarizes the types of variation on the protein level (Chong et al., 2015).

Table 2. Types of small protein altering variants.

Variant type	Transcript	Protein
Missense	**atAGA->atAGC**	**Arg->Ser**
Nonsense	**atAGA->atUGA**	**Arg->Stop**
Frameshift	**atAGA->atAAGA**	**Arg->Lys and frame shift**
Splice	**atAGA->acAGA**	**e.g. intron not spliced**

Changed nucleotides are highlighted, uppercase: exon and lowercase: intron.

Human diseases can be caused by acquired or genetic traits or by the combination of these two factors. For example, mononeuropathy can be caused by physical injury to a nerve resulting in non-progressive local symptoms, and hereditary polyneuropathy can be caused by a variant in a single gene leading to progressive neuronal degeneration and more spread symptoms (Rossor et al., 2013). When the disease is caused by a variant in one gene, it is called monogenic and is inherited through Mendelian inheritance (Figure 1). Most commonly a disorder is caused by a combination of hereditary and acquired traits where the cause and effect behind the disorder is the most difficult to study and pinpoint. For example, multiple predisposing variants in a number of different genes combined with unhealthy life choices can cause symptoms similar to hereditary neuropathy in some individuals (Liu et al., 2019).

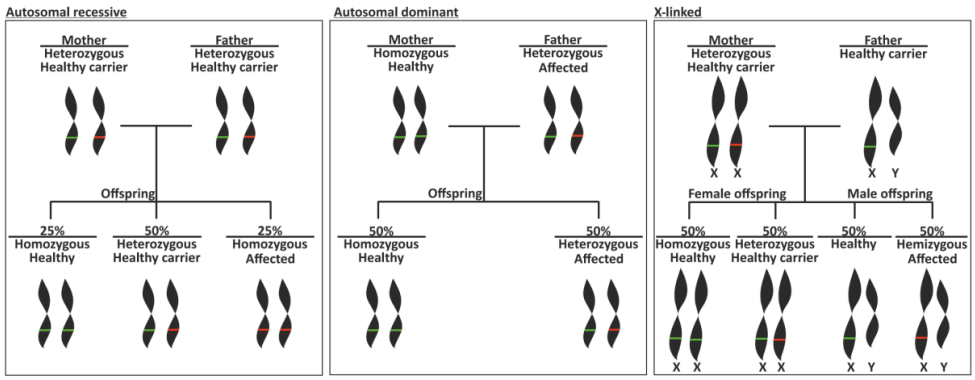

Figure 1. Mendelian inheritance in hereditary disease. In Mendelian inheritance a trait or a disorder is inherited in an autosomal recessive, autosomal dominant or X-linked fashion. Black shapes depict chromosomes, and the displayed percentages are likelihoods for the representative genotypes. Green indicates the wildtype allele and red indicates a mutated allele.

In monogenic Mendelian inheritance, trait determining variants are categorized as dominant and recessive. Dominant variants drive the phenotype even as heterozygous, one variant in one allele is enough. In the recessive case, both alleles must have a variant for the trait to be seen in the phenotype. In addition, hereditary diseases can arise *de novo* where the mutation

is absent in the parents but presents in the germline causing the offspring to have the variant and phenotype. When alleles reside in the X-chromosome, inheritance is termed X-linked. In X-linked recessive inheritance, males with only one X-chromosome are affected (Chong et al., 2015). In addition, 37 genes are inherited mitochondrially, of which 13 code for respiratory chain subunits. A variant in the mitochondrial genome (mtDNA) of the mother can be passed on to the child. In this case, the mutation load and threshold are crucial since mtDNA is multicopy genome. Each cell has hundreds to thousands of circular copies of its mtDNA, where some have the variant, and some do not. In homoplasmy all the copies carry the variant and in heteroplasmy only a portion have the variant. When a certain heteroplasmy level, threshold, of the mitochondrial genomes carries the variant, there is a phenotype. The threshold needed to cause symptoms varies between mutations and tissues (Filograna et al., 2021).

Most of the variants in the human genome are benign, thus they are not risk factors for a disorder and do not cause a hereditary disease. On the other extreme, some rare variants cause monogenic diseases where a single mutation in a single allele is enough to cause severe early-onset symptoms. Variants can be categorized with the American College of Medical Genetics (ACMG) guidelines into pathogenic, likely pathogenic and variants of unknown significance, depending on the level of evidence between the variant and the disorder (Table 3) (Richards et al., 2015). Briefly, pathogenic variants have been reported in multiple different families where the variant segregates with the disorder. In addition, the variant must be rare for recessive disease and not present in the healthy population for dominant disease, and the variant's effect on the protein product must have been studied on a molecular level. A variant is likely pathogenic when other similar variants in the same gene have been reported to be pathogenic, the variant is very rare in the healthy population, the variant segregates in the families with the phenotype, and there is molecular evidence of pathogeny on the gene level. When there is not enough evidence for the likely pathogenic category, but the variant is still of interest it can be called a variant of unknown significance (VUS).

Table 3. Clinical variant categorization.

Category	Simplified guideline	Example from thesis (I)
Pathogenic	**Variant previously associated with the phenotype in multiple families. Segregation in the family with the phenotype.**	***CHCHD10* heterozygous, c.197G>T, p.Gly66Val in adult-onset spinal muscular atrophy/Charcot-Marie-Tooth disease.**
Likely pathogenic	**New rare variant (<0.001) in a gene previously associated with the phenotype in multiple families. Segregation in the family with the phenotype. Functionally disruptive variant.**	***ABDC1* heterozygous, c.1885G>T, p. Asp629Tyr in adult-onset X-linked adrenoleukodystrophy.**
Variant of uncertain significance	**Rare variant (<0.001) in a gene previously associated with a similar phenotype in some cases. No supportive segregation data. Insufficient evidence of pathogeny.**	***COL6A2* heterozygous, c.728A>C, p.Cys1019Ser in adult-onset metabolic myopathy.**
Benign	**Common variant seen in healthy people in databases or seen in healthy family members. No previous disease association.**	**-**

Based on Richards et al., 2015.

On a molecular level, the mechanism of disease can be loss or gain of function, or haploinsufficiency. In gain of function, the variant introduces a novel pathogenic function or interaction into the mutant protein. The protein can for example gain an amylogenic element leading to its aberrant aggregation with interaction partners resulting in a neuropathy phenotype (Rebelo et al., 2016). In loss of function, the change in the protein causes reduction or complete loss of its original function in the cell. For example, *GAN* variants lead to reduced ability to regulate neurofilament (NF) levels, which also results in aggregation, filament accumulations and ultimately giant axonal neuropathy (Bomont et al., 2000). The net function of a protein can also be decreased by reduction of its amount or expression. In haploinsufficiency a missense variant, nonsense variant, deletion or regulatory area variant can cause an approximately 50% reduction in functional protein levels, which in some cases is sufficient to cause a disorder. As an example, reduction in the levels of HSPB8 by a frameshift variant causes myopathy (Echaniz-Laguna et al., 2017).

2.1.4 Genetic diagnosis

The genetic diagnosis, also called molecular diagnosis, of hereditary neuromuscular diseases has taken leaps forward thanks to developments in sequencing technologies. It is currently possible to sequence the whole human genome and find pathogenic variants from the sequencing data in just weeks.

2.1.4.1 Sanger sequencing

To unravel the genetic code of life different sequencing technologies were developed by multiple labs from the 1960s onward and a breakthrough was achieved by the Frederick Sanger lab in the mid-1970s termed as the chain termination technique (Sanger & Coulson,

1975). In the method, currently called Sanger sequencing, a DNA polymerase, a short nucleotide sequence complementary to the sequenced DNA (primer), the analyzed DNA sample, and normal as well as 3'hydroxyl group lacking radiolabeled nucleotides, dideoxynucleotides (ddNTPs), are mixed for the reactions. The ddNTP will terminate the reaction. Mixing four reactions, each with all nucleotides and low concentrations of a different radiolabeled ddNTPs, will generate DNA strands of all possible lengths. Running the DNA strands into four respective lanes in a gel, ddNTPs A, T, C or G containing reaction, autoradiography can be used to infer the genetic sequence from the relative location of the DNA bands on the gel. Because of its accuracy and ease Sanger sequencing has become the most widely used sequencing technology (Heather & Chain, 2016).

Current Sanger sequencing technologies use automated systems with fluorometric based ddNTPs allowing all four nucleotides to be detected in one reaction, and have replaced the use of gels by capillary based electrophoresis (Heather & Chain, 2016).

Sanger sequencing is an accurate and reliable method in discovering genetic variation, but it lacks in scalability. The sequence length (also called a "read") is restricted to approximately 1000 bases, and can be used to sequence a single genomic site with specific primers at a time. In the clinic Sanger sequencing is widely used when there is a strong indication of a certain genetic locus being responsible for the inherited disease. However, when symptoms can be caused by thousands of different variants in hundreds of loci, it is not feasible to use Sanger to sequence each of these sites one by one. When a positive variant has been found in a pedigree, Sanger sequencing is the optimal tool to investigate the inheritance pattern of the variant in the affected family (Heather & Chain, 2016).

2.1.4.2 Next-generation sequencing

Sanger sequencing dominated the field for decades before the development of next-generation sequencing (NGS) techniques. In the next widely used technology fluorescently labelled ddNTPs were not utilized, instead luminescence was measured after an enzymatic process with pyrophosphate, gaining the technology the name of pyrosequencing. Pyrophosphate, bound to a modified nucleotide, is released upon nucleotide addition to a growing polypeptide chain, and the pyrophosphate is then used as a substrate for ATP sulfurylase making ATP. The ATP is in turn used by a luciferase enzyme to produce a detectable signal. Different nucleotides are added sequentially and followed by washing or enzymatic degradation of the unbound nucleotides to get clear signal from a single base. DNA is locally amplified, before the sequencing reaction, to get strong enough signal for detection (Heather & Chain, 2016).

In the following widely used development, by Illumina, the nucleotides were again bound to fluorescent moieties. In the method a fluorophore at 3' hydroxyl position blocks the continuation of the polymerase reaction before it is cleaved. This allows the sequencing to happen in a synchronous manner, one nucleotide at a time. To get measurable signal the sequenced DNA sample is amplified on a chip, via bridge amplification to produce clonal

populations of DNA. In bridge amplification the single stranded DNA, with terminating complementary strands, is bound on an oligo covered surface and allowed to bend to bind other complementary priming oligo sites nearby. This results in local amplification of the DNA strand, which will be able to produce reliably measurable signal captured by a camera after fluorophore excitation (Heather & Chain, 2016).

The sequencing chip with locally amplified DNA strands is then sequentially bathed in modified nucleotides and excited producing a base specific fluorescent signal. Each nucleotide bath is followed by enzymatic cleavage of the fluorophore to allow the polymerase to continue in the next cycle. The read length produced by the Illumina system is relatively short (35bp), but the bridge amplification allows the sequencing to be redone from the other direction to achieve paired-end reads. Thus, both ends of the DNA segments will be sequenced allowing more precise alignment to the genome. With the parallel sequencing of millions of reads simultaneously it has become feasible to sequence large areas or even the whole human genome to identify pathogenic variants. The sequenced reads are mapped to the human genome and compared to the reference sequence to find variants in the sequenced individual (Gorcenco et al., 2020; Heather & Chain, 2016).

For the clinic NGS technology has been utilized to develop specific gene panel sequencing platforms to cost-efficiently sequence tens to hundreds of genes at the same time. These panels capture sequences of certain set of genes, such as neuropathy associated loci (Ylikallio et al., 2014). The targeted regions typically have a good coverage, and the data is easy to analyze, because of its relatively small size. Coverage in panels is fairly uniform between samples allowing the detection of exonic deletions and duplications. Patients analyzed at the same time can be used as reference coverage to exon regions (Gorcenco et al., 2020).

The drawback of NGS panels is their relatively small coverage of the genome. Only the selected genes are sequenced, and the selection is based on historically determined interesting genes. A pathogenic variant can reside in a gene not covered by the panel. In addition, the clinician must order a specific panel, which might not be the right panel for the patient and then sequentially order additional panels if no variants are found (Gorcenco et al., 2020).

The power of current NGS technologies is best served in whole exome (WES) or genome sequencing (WGS). In WES probes are used to capture all coding loci, exons (1-2% of the whole genome, and approximately 96% of all protein coding regions) (Gorcenco et al., 2020; Pipis et al., 2019), and in WGS the whole genome is sequenced. Therefore, WES can be used to find variants in all exons and WGS in all genes including introns and all intergenic regions. The amount of data produced by the NGS methods is huge where WES can capture up to tens of thousands of variants and WGS up to millions. The clinical utility of different sequencing techniques is summarized in Table 4.

Table 4. Sequencing methods used in the clinic.

Technology	Genomic coverage	Read quantity	Optimal for	Variants (amount)	Data processing load	Price	Clinical usability	Drawbacks
Sanger	**Single gene**	**Very high**	**Small and indel variants**	**Few**	**Small**	**Low**	**Situational**	**Low scalability**
Panel	**Tens to hundreds of genes**	**High**	**Small, indel and copy number**	**Hundreds**	**Small**	**Low**	**Good**	**Biased selection of genes**
WES	**All protein coding regions (1-2%)**	**High but uneven**	**Small, indel and potentially copy number**	**~100 000**	**Moderate**	**Moderate**	**Good**	**Uneven coverage**
WGS	**Whole genome (95-98%)**	**Moderate but even**	**Small, indel and copy number**	**~5million**	**Heavy**	**High**	**Laboursome**	**Variant filtering**

Based on Gorcenco et al., 2020.

2.1.4.3 Variant interpretation

The huge amount of variant data produced by WES or WGS must be filtered to find possible pathogenic variants among the benign. Most of the variants are common SNPs, which do not cause monogenic disease. In addition, many of the variants are in regions with no previous association with human disease. The ACMG guidelines (Richards et al., 2015) provide a framework for categorizing variants from NGS data (Table 5). Variant or allele frequency in the population is the first step to determine the pathogeny of a variant, since variants causing monogenic disorders are rare in general population. The frequency of a variant in a population can be determined from public databases, such as GnomAd v2.1.1, which contains 125 748 exomes and 15 708 genomes (Karczewski et al., 2020). In general terms a variant causing a dominant severe early-onset disorder should not be present in the general population, but recessive variants can be found in the general population with a high carrier frequency. For example variants in the *SLC17A5* gene causing recessive Salla-disease can be found in 1 per 200 in Finland (Aula et al., 2000). Therefore, for determining an acceptable variant frequency it is important to define the likely mode of inheritance of the hereditary disease: dominant, recessive, X-linked or mitochondrial.

Table 5. Variant filtering tools

Method	Tool	Info	Reasoning
Frequency	**GnomAd**	**Collection of WES and WGS data**	**Pathogenic variants are rare in the healthy population**
Variant annotation	**NCBI genome assembly**	**Genome reference and variant annotation**	**Genetic variant location and if in a gene consequence on transcript and protein level**
Segregation	**Sanger sequencing**	**Cheap and efficient sequencing of single variants**	**Confirmation of a variant and segregation in the family. Correlation with variant and phenotype.**
In-silico prediction	**PolyPhen-2**	**Protein function and structure algorithm**	**Used to predict functional and structural changes in the protein caused by the variant.**
Disease association	**OMIM**	**Database of variant disease association**	**Previous variant associations with disease. Genotype and phenotype correlation.**

Rarity of a variant does not determine its pathogeny. All humans carry rare variants in their genome, but the majority of these are benign. Variants outside genes or in the heterochromatin are less likely to have an impact and cause disease. In contrast, variants in genes and especially in exons are more likely causes of disease. Not only the genetic location but also the variant's effect on the gene predicts its pathogeny, ~85% of pathogenic variants reside in the exons and cause a change in the translated protein (Gorcenco et al., 2020). Variants that change the amino acid (missense), reading frame (insertion or deletion of nucleotides not dividable by three), stop-codon or splice site potentially lead to severe protein level alterations. Also, variants in essential regulatory elements and promoters can change the gene expression level or pattern and cause symptoms. Therefore, variants altering the protein sequence or amount of protein or mRNA should be regarded as the most likely to cause disease (Richards et al., 2015).

Additionally, gene functional studies and *in silico* prediction tools are used to assay the possible pathogenicity of variants. *In vitro* functional studies can be used to assess the protein function and the variants' effects. For example, a single amino acid change can result in protein's loss-of-function through degradation or in toxic gain-of-function via aggregation. A less laborsome approach to variant assessment is the use of *in silico* predictive algorithms. They can aid in the categorizing of variants but lack experimental power. The algorithms assay evolutionary conservation (i.e. SIFT), protein function and structure (i.e. PolyPhen), and variant frequency (i.e. GnomAd z-score). The computational

data can be used to guide variant annotation but cannot be used as a sole predictor of pathogenicity (Richards et al., 2015).

Variant association with disease can be interrogated from disease databases. Online Mendelian inheritance in man (OMIM, https://omim.org/) catalogues clinical and genetic data from publications, to link phenotypes to genotypes. OMIM mainly contains data on pathogenic and likely pathogenic variants. ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/), on the other hand, gathers also unpublished genetic and phenotype data, as well as VUS and benign variants. Association of mitochondrial variants with disease is collected into the Foswiki MitoMap (https://mitomap.org/MITOMAP) database. The data deposited to the above-mentioned databases is not completely curated making it imperative to use additional methods to determine a variant's previous association with disease (Richards et al., 2015; Shah et al., 2018).

Previously curated pathogenic variants are straightforward to confirm in patients. A clinical overview and additional Sanger sequencing of the variant in the patient and immediate family is typically enough to determine the molecular diagnosis. In the case of new genetic variants in known disease genes, a more complete genetic and clinical evaluation is needed. Methods described above can be used to assess the pathogenicity of the new variant. Gene level clinical data, i.e. from OMIM, can be used to compare the new variant to previous cases with pathogenic variants in the same gene. Additionally, Sanger sequencing of the variant in the family of the patient coupled with family members' clinical data is an effective tool in determining variant pathogenicity through segregation.

Initial analysis of genetic data can result in no potential findings. Nonetheless, the filtering can be repeated periodically to uncover variants determined as pathogenic after the initial analysis as in (Sainio et al., 2019). Additionally, variant and pedigree data on uncertain variants can be shared on websites such as Genematcher (https://genematcher.org/) to locate patients with similar variants and phenotypes to discover new pathogenic genes (Rönkkö et al., 2020; Ylikallio et al., 2017).

Since WES can miss duplications or deletions of large areas and WES and WGS are not optimal in sequencing large repeat-elements, NGS can be complemented with alternative methods to catch large chromosomal alterations, repeat expansions and mitochondrial variants.

2.2 Hereditary neuromuscular disorders

Hereditary neuromuscular disorders are clinically and genetically highly diverse. For example, mutations in over a hundred genes have been found to cause CMT (Klein, 2020) and more than 30 have been associated with LGMD-like phenotypes (Straub et al., 2018). A common feature in most cases is muscle weakness, but the primary affected tissue (Figure

2) and pathological molecular mechanism leading to the disorder depend on the affected gene. Therefore, discovering the genetic etiology behind the symptoms is paramount in understanding the molecular mechanism of disease. Understanding the molecular pathways of the diseases will possibly enable the development of disease altering drugs. Moreover, the grouping of diseases caused by different genes that converge on the same pathological pathways could facilitate the development of treatments for larger patient groups.

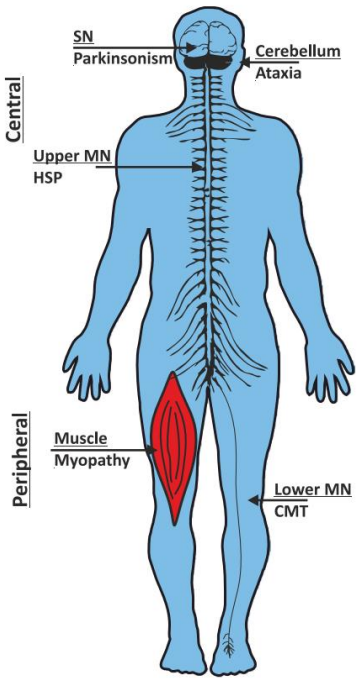

Figure 2. Tissues affected in hereditary neuromuscular and neurological disease affecting movement. Representative schematic of primary affected tissue in different diseases studied in this thesis. Central nervous system neurons are affected in HSP, ataxia and Parkinsonism, and peripheral tissues in CMT and myopathy. Hereditary spastic paraplegia (HSP), Charcot-Marie-Tooth disease (CMT), motor neuron (MN) and substantia nigra (SN).

2.2.1 Limb-girdle muscular dystrophy

LGMD is part of hereditary myopathies, a larger group of disorders caused by mutations that disrupt the structure and/or function of muscular proteins. Degenerative muscle diseases caused by variants in muscular proteins are typically called myopathies or muscular dystrophies. Disorders that affect the neuromuscular junction are termed myasthenic syndromes. Myopathies are further categorized upon the disease gene or gene groups that typically cause similar symptoms. LGMD is a general term for muscle-oriented progressive limb and proximal girdle muscle weakness and wasting that is not congenital or X-linked. The disease can manifest in early childhood or late-adulthood, and the severity varies greatly. Some patients experience mild difficulties in walking with no dystrophy observed in muscle biopsies, and some become wheelchair-bound during the course of the disease with severe dystrophic changes in musculature. LGMD is often restricted to skeletal muscle, but in some cases respiratory or cardiac muscles are affected (Dowling et al., 2021).

The complexity of neuromuscular disorders can be further appreciated with a recently discovered disease gene Pyridine Nucleotide-Disulphide Oxidoreductase Domain 1 (*PYROXD1)* causing varied myopathic symptoms. The disease was initially described as an early-onset myopathy with mixed features (O'Grady et al., 2016). Later reports described large phenotypic variability, ranging from mild elderly-onset LGMD to congenital myopathy leading to the loss of ambulation. Furthermore, some patients showed additional signs of neuropathy (Daimagüler et al., 2021).

2.2.1.1 Limb-girdle muscular dystrophy genetics

LGMD is caused by recessive (LGMDR) and dominant (LGMDD) mutations in a group of over 30 genes (Table 6). Common disease genes for LGMD worldwide include *Sarcoglycan Gamma (SGCG), Sarcoglycan Alpha (SGCA), Dysferlin (DYSF), Calpain 3 (CAPN3), Fukutin Related Protein (FKRP)* and *Anoctamin 5 (ANO5)* (Liewluck & Milone, 2018). Causative genes vary in prevalence in different populations with *DYSF* being more common in Southern Europe (Georganopoulou et al., 2021) and *ANO5* in Northern Europe (Hicks et al., 2011; Penttila et al., 2012). LGMD has genetic and clinical overlap with other myopathies, such as distal myopathies and congenital muscular dystrophies. For example, mutations in *ANO5* that disrupt muscle sarcolemma integrity can cause anoctaminopathy-5 (LGMD2L) and Mioshi myopathy. In addition, *ANO5* mutations further contribute to the phenotype variability inside the LGMD category, because the patients might display asymmetric atrophy of thigh and upper arm musculature not seen with other LGMD genes (Bolduc et al., 2010).

Table 6. LGMD genes common in Finland.

Omim	Gene	Inheritance	Subgroup	Other phenotypes
LGMDR1/D4	***CAPN3***	**AR/AD**	**Calpainopathy**	**-**
LGMDR2	***DYSF***	**AR**	**Muscular dystrophies with defective membrane repair**	**AR Miyoshi muscular dystrophy 1 & AR Myopathy, distal, with anterior tibial onset**
LGMDR3-6	***SGCA,B, C, D***	**AR**	**Sarcoglycanopathies**	**-**
LGMDR9	***FKRP***	**AR**	**α-dystroglycanopathies**	**AR, congenital Muscular dystrophy-dystroglycanopathy**
LGMDR10	***TTN***	**AR**	**Z-disk proteinopathies**	**AD cardiomyopathy, AD myofibrillar myopathy, AR Salih myopathy**
LGMDR12	***ANO5***	**AR**	**Muscular dystrophies with defective membrane repair**	**AR Miyoshi muscular dystrophy 3, AD Gnathodiaphyseal dysplasia**
LGMDR19	***GMPPB***	**AR**	**α-dystroglycanopathies**	**AR congenital Muscular dystrophy-dystroglycanopathy**
LGMDD1	***DNAJB6***	**AD**	**Z-disk proteinopathies**	**-**

Based on Straub et al., 2018.

2.2.1.2 Mechanisms of limb-girdle muscular dystrophy

LGMDs are often caused by variants in genes involved in the dystrophin glycoprotein complex (DG), which resides in the myofiber sarcolemma. Such genes include *β sarcoglycan* (*SGCB)*, which is a transmembrane sarcoglycan in the sarcolycan subcomplex part of the DG. The DG complex is suggested to bridge the binding of the actin cytoskeleton and the extracellular matrix. Within the DG complex the sarcoglycans mechanically stabilize the myofiber sarcolemma and transduce mechanical information. Current research indicates variants in the sarcoglycans to reduce the number of functional sarcoglycan complexes. This in turn destabilizes the DG making it susceptible to damage during muscle contraction. LGMD causing genes do not only code proteins of the DG complex, but they also function in the sarcomere, nuclear envelope as well as the muscular triad (Ozawa et al., 2005; Tarakci & Berger, 2016).

All in all, LGMDs are clinically an enigmatic group of disorders with varied ages of onset and progression patterns. Unfortunately, it can be difficult to discern hereditary LGMD from other causes of muscle weakness. Genetic methods are a widely used tool to help clinicians to achieve a molecular diagnosis for patients with difficult to diagnose myopathic disorders. NGS-panel with the 31 most common LGMD genes resulted in a genetic diagnosis for 34% of LGMD patients (Özyilmaz et al., 2019). Additionally, a genetic diagnosis is needed to further study the molecular mechanisms of disease (Dowling et al., 2021).

2.2.2 Charcot-Marie-Tooth neuropathy

Motor symptoms originating from peripheral neuron degeneration or disruption can be caused by Charcot-Marie-Tooth (CMT) disease, and hereditary motor neuropathy (HMN). Additionally, predominantly sensory neuropathy is termed hereditary sensory autonomic neuropathy (HSAN) (Klein, 2020). For readability and because the genetic and clinical features and nomenclature of CMT, HMSN and HMN overlap (Beijer & Baets, 2020), I will discuss these disorders collectively as CMT. CMT is often referred to as the most common hereditary neurological disease with an incidence of 1 in 2500 (Skre, 1974). In Northern Ostrobothnia of Finland the prevalence of CMT is estimated to be 1 in 2800 (Marttila et al., 2017). CMT is divided into demyelinating CMT1 and axonal CMT2 disease categories. In a simplistic view, genes causing a disease predominantly with a dominant inheritance and involved in the myelination or maintenance of myelin are mutated in CMT1, causing reduced (<38m/s) nerve conduction velocity (NCV). In CMT2, variants in genes cause the disorder with dominant and recessive inheritance. The genes can be involved in varied axonal processes causing neuron degeneration within the neuron when mutated. In contrast to CMT1 NCV is rarely affected (NCV>38m/s) in CMT2 but compound muscle axon potentials (CMAP) can be reduced. In practice, the clinical distinction between CMT1 and 2 is not always clear, the symptoms are similar and additional studies are needed to assess the presence of demyelination. In some reports intermediate CMT is used to describe CMT with axonal and demyelinating features, often with moderately reduced NCV (25 m/s-45

m/s). CMT can involve motor and sensory symptoms, therefore HMN is sometimes used to describe neuropathy purely restricted to motor neurons (Klein, 2020; Rossor et al., 2016).

Clinically CMT causes length dependent muscle weakness, atrophy, and sometimes sensory symptoms. The symptoms can start from birth to elderly age, and normally progress slowly during life. Severity and progression of disease varies significantly, from early loss of ambulation to difficulties in rising stairs in elderly years. Symptoms tend to arise in a length depended manner, where the longest axons of the body are affected first. This causes initial weakness in feet muscles and loss of sensation in the toes. Loss of neural connections to the feet leads to a common feature of CMT - pes cavus. In pes cavus foot muscles are severely atrophied causing tightening of tendons and muscles resulting in an arched foot (Klein, 2020).

Clinically CMT and other neuropathies have overlapping features with ALS, spinal muscular atrophy, hereditary spastic paraplegia and other motor neuron disorders emphasizing the need for genetic diagnosis (Beijer & Baets, 2020; Klein, 2020).

2.2.2.1 Charcot-Marie-Tooth neuropathy genetics

Mutations in over one hundred genes have been reported to cause CMT (Table 7). More than 60% of all genetically diagnosed CMT cases, including axonal and demyelinating, are caused by a 1.4 Mb duplication in chromosome 17p containing the *peripheral myelin protein 22* (*PMP22)* gene (Pipis et al., 2019). The autosomal dominant duplication causes a childhood-onset demyelinating neuropathy leading to secondary neuronal degeneration. Interestingly, the deletion of the same area in chromosome 17 containing *PMP22* causes hereditary neuropathy with liability to pressure palsies. After PMP22, the most common causes of CMT1 are dominant mutations in *myelin protein zero* (*MPZ)* resulting in a very similar phenotype as with *PMP22* duplication, and X-linked dominant variants in *gap junction protein beta 1* (*GJB1)* resulting in adolescent-onset neuropathy occasionally companied with central symptoms. CMT2 is caused by a myriad of different genes and mutations, where no gene clearly predominates as in CMT1. Nonetheless, variants in *mitofusin 2* (*MFN2*) are the most likely causes of CMT2 in approximately 20% of patients worldwide. However, in the Finnish population *GDAP1* variants are more prevalent (Marttila et al., 2017). Recessive variants in *MFN2* cause a severe axonal neuropathy typically leading to loss of ambulation before the age of 20. Overall genetic yield in hereditary neuropathies varies, with only approximately half of clinically categorized CMT patients receiving a molecular diagnosis. In CMT1 the diagnostic yield is up to 85%, but in CMT2 and HMN diagnostic yield is between 15-40% (Klein, 2020; Murphy et al., 2012; Pipis et al., 2019).

Table 7. Common CMT genes.

Omim	Gene	Inheritance	Subgroup	Other phenotypes
CMT1A/E	***PMP22***	**AD**	**Demyelinating**	**AD neuropathy with pressure palsies, AD/AR Dejerine-Sottas disease, AD Roussy-Levy syndrome**
CMT1B/2J/int D	***MPZ***	**AD**	**Demyelinating**	**AD/AR Dejerine-Sottas disease, AD Hypomyelinating neuropathy, AD Roussy-Levy syndrome**
CMTX1	***GJB1***	**XLD**	**Demyelinating**	**-**
CMT2A2A/2A2B	***MFN2***	**AD/AR**	**Axonal**	**AD Hereditary motor and sensory neuropathy VIA**
CMT2K/4A/int A	***GDAP1***	**AD/AR**	**Axonal***	**AR Charcot-Marie-Tooth disease with vocal cord paresis**
CMT1F/2E/int G	***NEFL***	**AD/AR**	**Axonal***	**-**

*Reduction in NCV, intermediate = int. Based on Klein 2020 and Bomont 2021.

2.2.2.2 Mechanisms of neuropathy

The molecular mechanism of neuropathy depends on the mutated gene and even the specific mutation in the gene (Figure 3). In CMT1, where myelin related genes are often mutated, the disease is a consequence of disrupted myelin formation or maintenance by Schwann cells. For example, the duplication of *PMP22* causes demyelination with a toxic gain of function mechanism (Li et al., 2013; Watila & Balarabe, 2015). The increased expression of *PMP22* is thought to disrupt the growth and differentiation of Schwann cells (Shi et al., 2018) and inhibit proteasome activity and autophagy (Fortun et al., 2006) leading to reduced functional myelin. Nerves in patients and animal models display myelin loss and onion bulb like formations, formed from proliferating Schwann cells attempting to remyelinate the axons (Hanemann et al., 1997; Sereda et al., 1996).

In axonal CMT2 mainly neuronal homeostasis processes are affected. The disease mechanism of most CMT2 mutations is still unclear. Nevertheless, in one of the most common causes of CMT2, mutations in *MFN2* cause mitochondrial trafficking defects (Mou et al., 2021), tubulin acetylation disruption (Picci et al., 2020) and mitochondria-ER communication impairment (Bernard-Marissal et al., 2019). Thus, multiple mechanisms of disease have been suggested. Most likely axonal degeneration in association with *MFN2* mutations, and in CMT overall, is caused by a combinatory effect of multiple disrupted pathways.

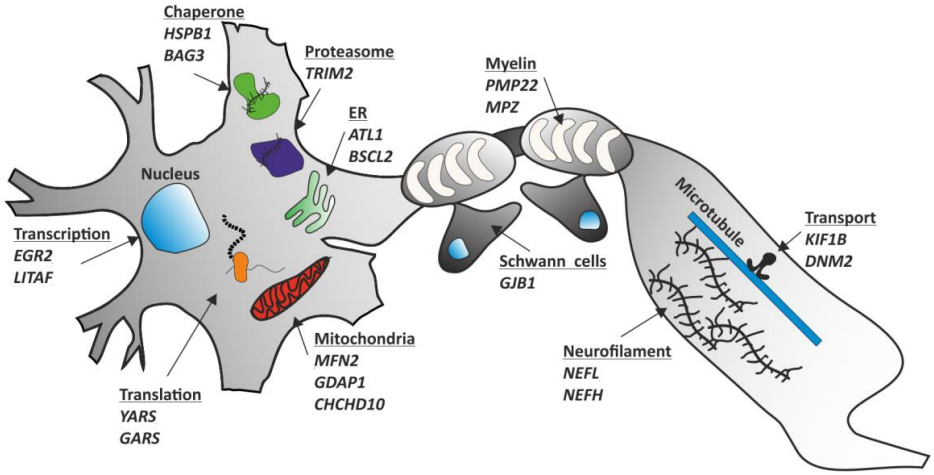

Figure 3. Genes and cellular complexes affected in CMT. Representative set of genes causing CMT by affecting various cellular processes. Variants in all the genes cause peripheral axon degeneration, either directly or via demyelination. Based on Klein et al., 2020.

2.3 Neurofilament light and neurofilaments

Axonal (CMT2E) and demyelinating CMT (CMT1F) are caused by dominant and recessive mutations in *NEFL,* respectively. NFL is a type IV intermediate filament, which forms NFs in neurons with other neuronal intermediate filaments neurofilament medium (NFM), NFH, α-internexin (INA) and peripherin (PRPH). NFs fill the axoplasm and regulate axonal caliber (Bomont, 2021). Therefore, NFs have long been thought to have a quiescent and stable structural role in the neurons. In contrary to actins and tubulins, intermediate filaments are non-polar structures not actively polymerizing and depolymerization. However, NF subunits are transported along the axons and can polymerize into long structures (Wang et al., 2000; Yuan et al., 2009). Furthermore, recent studies have elaborated more functional and dynamic roles for NFs.

2.3.1 Neurofilament formation

NFs are neuronal intermediate filaments (IF). The formation of long IF structures follows the same steps in all cells with cell type specific IF proteins. IF formation has been extensively studied with muscle desmin and skin vimentin. NFs, as other IFs, are formed of IF protein oligomers organized into 10nm thick elongated structures (Herrmann & Aebi, 2016). IF proteins are structurally organized into five domains: three conserved central core domains; Coil 1a, 1b and 2, one short N-terminal head domain and a variable length C-terminal tail domain (Figure 4). The central core domains fold into a-helices, which coil around another filament to form parallel dimers (Geisler & Weber, 1982). To assemble IFs, two parallel homo- or heterodimers bind each other with their N-terminal coils (Coil 1) to form a tetramer (Chernyatina et al., 2012). Intermediate filament tetramers are the basic

building blocks for further filament organization (Chernyatina et al., 2015). Eight tetramers form unit-length filaments (UFL) via lateral association, the ULFs being the building blocks of mature IFs (Herrmann et al., 1996). The mature elongated filament is built from longitudinal annealing of ULFs and a final compaction of the filament produces the 10nm diameter IF (Herrmann & Aebi, 2016).

In NF formation the small molecular weight neuronal intermediate filaments, NFL (68kDa), INA (~60kDa) and PRPH (~60kDa), form the central core dimer of the shaping mature NF. Dimers are also formed by combination of the short and longer NF proteins, NFM (~160kDa) and NFH (~200kDa), which have long C-terminal tails that protrude out of the protofilament. Dimers formed from the small molecular weight NFs are needed for tetramer and ULF formation, therefore dimers of NFM or NFH cannot solely assemble into mature NFs (Gafson et al., 2020). Structural images of NFs have revealed a bottlebrush-like molecular structure where the protruding NFM and NFH tail domains from tetramers resemble the brushes of a bottlebrush (Hisanaga & Hirokawa, 1988). Composition of NFs differs between central and peripheral neurons. Central neuron NFs are mainly formed of NFL, NFM and NFH as well as INA (Yuan et al., 2006). In peripheral neuron NFs INA is exchanged to PRPH (Yuan et al., 2012). Early in neuronal progenitor development IF dimers used for tetramers and ultimately ULF can be formed from vimentin and nestin. Later in neuronal differentiation and axon development vimentin and nestin will be replaced with NFL and INA or PRPH. During further axonal elongation and caliber growth also NFM and NFH are used to form rigid long lasting NFs. Furthermore, NFM and NFH tails are heavily phosphorylated to stabilize the IF network. Mature neurons do not express vimentin or nestin (Nixon & Shea, 1992).

NFs are transported along the axons as heterodimers and short assembled filaments, which are incorporated into the NF network (Wang et al., 2000; Yuan et al., 2009). Motor proteins such as Kinesin-V regulate NF transport (Lina Wang & Brown, 2010). The short filaments have been shown by photobleaching experiments to anneal end-to-end to the growing mature NF (Uchida et al., 2013). Thus, NFs exist in the axons as transported dynamic short filaments and long stationary filaments part of the axon cytoplasm. Mature neurons with an established network of NFs mainly consist of stationary NFs maintained by a small dynamic population of short NF oligomers (Yuan et al., 2009). Additionally, at least *NEFL* mRNA is anterogradely transported in transport granules along the axon to be locally translated, and therefore likely oligomerized and incorporated into the NF network at distal neuronal compartments (Alami et al., 2014; Krichevsky & Kosik, 2001).

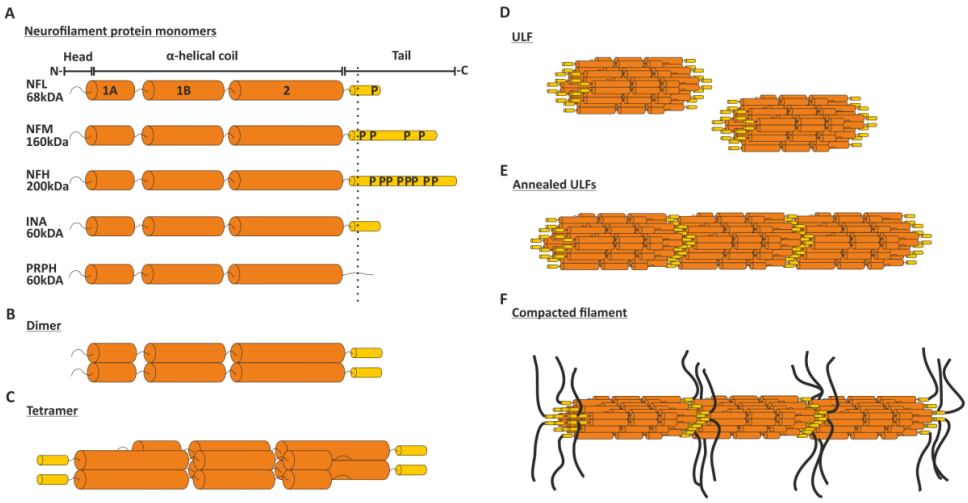

Figure 4. Neurofilaments and their assembly. A. Neurofilament (NF) protein size and domains. Dashed line depicts Tail domains to be longer than in the figure. P represents phosphorylation sites. B. NF dimers are formed of parallel hetero- or homodimers. C. NF tetramers are formed of two antiparallel dimers. D. Unit-length-filaments (ULF) are formed from eight tetramers. E. ULF anneal end-to-end to form long filamentous structures. F. Mature NFs compact to generate 10nm diameter rigid filaments. NFH and NFM tails protrude from the filament core resulting in a bottlebrush-like appearance. Based on Bomont et al., 2021.

2.3.2 Neurofilament modification

NF proteins are post-translationally modified by phosphorylation, glucosamination, nitration, oxidation and ubiquitination. Serine and threonine moiety phosphorylation, by multiple different kinases such as cyclin-dependent kinase, protein kinase C, and calcium/calmodulin-dependent protein kinase type II, is the major modification of all NF proteins. All N-terminal heads of NF monomers can be phosphorylated, but only C-terminal tails of NFM and NFH are extensively phosphorylated (Nixon & Sihag, 1991). The phosphorylation of head domains mainly occurs in the cytoplasm and of C-terminal tail domain in the axons during radial growth and myelination (Nixon et al., 1994). The extensively phosphorylated tails of NFM and NFH are thought to protect the NF proteins from degradation (Rao et al., 2012). NFs have been shown *in vivo* to have unexpectedly long half-lives, up to 8.5 months in mice (Millecamps et al., 2007). NF ubiquitination has been shown to regulate the degradation of NFs in giant axonal neuropathy (GAN) (Bomont et al., 2000) and in CMT caused by mutations in *TRIM2*, (Ylikallio et al., 2013), in which mutations in NF ubiquitinating proteins cause NF aggregates and axonal neuropathy. In addition to ubiquitin mediated degradation, the chaperone Heat Shock Protein Family B (Small) Member 1 (HSPB1) is involved in NF subunit degradation (Haidar et al., 2019) and the sacsin protein complex can block NF assembly (Gentil et al., 2019).

2.3.3 Neurofilament functions

NFs provide structural integrity and protect long axons, such as lower motor neuron axons, from mechanical stress (Bomont, 2021; Herrmann et al., 2009). NFs are speculated to be the most significant load-bearing component of neurites (Grevesse et al., 2015). Nevertheless, NF function in neurons is not restricted to mechanical stress protection by the viscoelastic properties of the NF network (Figure 5). Similarly to microtubules and spectrins along the axons, the NF network works as a scaffold for interactions with neuronal proteins and organelles (Yuan et al., 2017). By binding microtubule monomers and stathmin complexes NFs regulate microtubule assembly (Bocquet et al., 2009; Yadav et al., 2016). Therefore, axons with more NFs tend to have fewer microtubules and vice versa. In addition, single NF proteins have separate functions in synapses that indicate a regulatory role for NF in neurotransmission. In the post-synapse, NFL can bind ionotropic glutamate NMDA 1 receptor subunits (GRIN1) in multimeric glutamatergic receptors and stabilize them on the membrane. The loss of NFL in mouse hippocampal neurons showed a reduction in GRIN1 density and reduced excitatory post-synaptic potentials (Yuan et al., 2018). Furthermore, NFM regulates dopamine receptor anchoring by binding to endosomes carrying the receptor (Yuan et al., 2015).

NFs in the axonal cytoskeleton form a network with docking sites for reversible interactions with molecular motors and vesicular organelles (Yuan et al., 2017). Furthermore, NFL anchors mitochondria, ER, endosomes and lysosomes through Myosin-Va interaction and direct binding with phospholipids, and thus regulates vesicular and organelle distribution in the axons (Kim et al., 2011; Rao et al., 2002, 2011). Also, direct NF and mitochondria binding has been suggested in a cell free setup (Wagner et al., 2003). *NEFL* KO studies have demonstrated that NFL regulates mitochondrial movement. Increased mitochondrial movement was reported in dorsal root ganglia lacking NFL protein (Perrot & Julien, 2009), whereas aggregating NFL mutants have shown decreased movement of mitochondria (Gentil et al., 2012; Saporta et al., 2015). Further interactions have been reported for INA and PRPH, which bind adaptor protein-3 (AP-3). With the AP-3 interaction, NFs are speculated to regulate late-endosome and lysosome distribution and maturation (Styers et al., 2004).

NFs are responsible for axonal caliber growth and maintenance of myelinated axons. Initial studies on *NEFL* KO mice (Zhu et al., 1997) and Japanese quails with a nonsense mutation in *NEFL* (O. Ohara et al., 1993) revealed the relationship with NF quantity in myelinated axons and axonal size. Loss of NFL or NFM (Jacomy et al., 1999) reduces the amount of NFs in myelinated axons and leads to a decrease in axon caliber. The axon caliber regulates conduction velocity, with larger caliber axons having faster conduction. Therefore, NFs regulate neuronal signal conduction velocity in myelinated axons via axonal caliber modulation. Furthermore, NFH has been shown to directly modulate ion-channel function. Loss of NFH reduced nerve conduction velocity without affecting the axon caliber (Križ et al., 2000). However, axon caliber and conduction velocity are reduced in the combined

deletion of NFM and NFH tail domains, possible because of the loss of the highly phosphorylated regulatory moieties (Garcia et al., 2003).

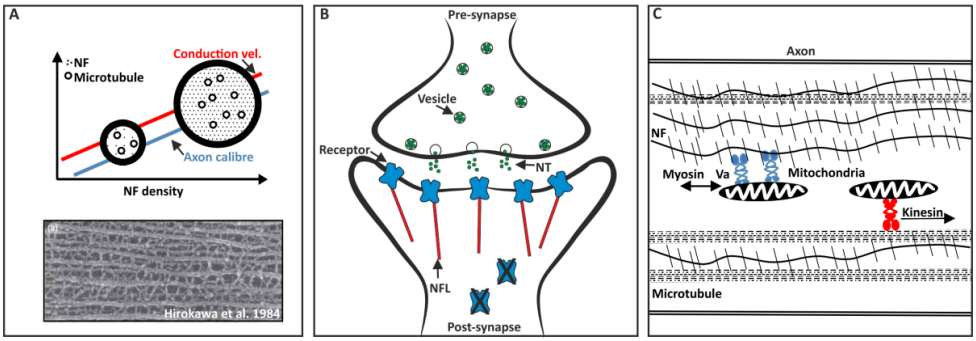

Figure 5. Functions of neurofilaments. A. Neurofilaments (NFs) fill the axoplasm and support axon caliber growth. Large diameter axons have more NFs and faster conduction velocity than small caliber axons. NF neurofilament. B. NFL binds to glutamatergic receptors and stabilizes them on the post-synaptic membrane reducing their degradation. NT neurotransmitter. C. NFs function as scaffolds for stationary organelles in the axon via Myosin Va. Based on Bomont 2021 and Hirogawa et al., 1984.

2.3.4 *NEFL* mutations in CMT

Over 30 different dominant and recessive mutations in *NEFL* cause CMT1F/2E (Figure 6) (Stone et al., 2021). The mutations are distributed throughout the gene and most are present in a single family or patient. Dominant missense variants are the most common, and variants p.P8R, p.P22S, p.N98S and p.E396K have been found in six or more sporadic cases/families. Interestingly, missense variants are aggregated in the beginning of the head domain, and head-rod and rod-tail junctions. In genetically studied neuropathy cohorts approximately 1% of patients' disorders are caused by *NEFL* variants (Murphy et al., 2012; Stone et al., 2021).

CMT caused by *NEFL* variants is clinically diverse. Assessing the phenotypic effect of variant location in the protein is difficult since the most common variants dominate the data and many of the variants reside around the head-rod (~84aa) and rod-tail (~393aa) junctions. Nevertheless, some indications are discernable from data gathered in Stone et al., 2021 (Stone et al., 2021) from all published (39 publications) *NEFL* variants totaling to 168 cases. Age of disease onset varies from childhood to late adulthood. For example, the most common dominant variants listed in previous paragraph result in disease onset from 0.3 to 55 years of age. On average the disease onset of *NEFL*-linked CMT (missense and nonsense pooled) in different protein domains is 9.8 years in head domain, 15.0 years in rod and 13.6 years in tail. Nevertheless, nonsense variants in the rod domain and missense variants in the head domain and head-rod junction tend to cause symptoms to arise earlier. Complementary to earlier onset, the same variants are more likely to display a severe reduction in NCV and reduction in CMAP. Missense rod and tail domain variants have slightly reduced or normal

NCV and CMAP in the normal range. The overall severity of the disease is not affected by variant location with a CMT neuropathy score analysis of a small group of cases (37/168 patients) (Stone et al., 2021). In addition to variability between different variants, there is high heterogeneity in disease severity with the same variant, even in the same family (Boerkoel et al., 2002).

The spread of variants along the entire protein indicates that the disrupted function of a single protein domain or structure is not responsible for the disease. Furthermore, the dominant toxic gain-of-function mechanism cannot be the common mechanism for all *NEFL*-linked CMT if nonsense mutations lead to loss of protein but result in the same disease with increased severity. Nevertheless, a common feature seen in humans and animal models is the loss of NFs or their correct alignment in myelinated axons in *NEFL* missense and KO mice. A deeper understanding of the functions of oligomeric NFs and their monomers is needed to unravel the disease mechanism in *NEFL* CMT. Surprisingly, even though nonsense mutations in *NEFL* cause an early-onset disease, deletion of the *NEFL* gene has been used in disease models to ameliorate neurological symptoms. For example, deletion of *Nefl* in KO mice of tubulin binding factor E (*pmn*) restored axonal transport and ameliorated motor symptoms. *Pmn* mice are used to model motor neuron diseases as they present with reduced microtubules and axonal transport, and the dying back of peripheral axons. The deletion of *Nefl* in these mice increases the amount of tubulins possibly through reduced NF regulation through stathmin complexes, resulting in milder motor symptoms (Yadav et al., 2016).

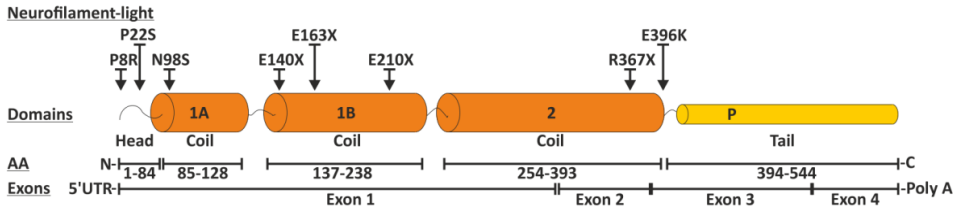

Figure 6. Variants in *NEFL* causing CMT. The most common dominant and all nonsense variants in *NEFL* causing CMT, and their amino acid change and position. Variant locations are shown in protein domain, amino acid (AA) and transcript levels. Based on Stone et al., 2019 and Stone et al., 2021.

2.3.5 *NEFH* mutations in CMT

As NFs have multiple important functions in neurons, it is not surprising that mutations causing CMT are also found in another NF gene: *NEFH*. Heterozygous frameshift stop-loss mutations in *NEFH* lead to the translation of cryptic amylogenic elements from the 3'-UTR of the mRNA after the canonical stop-codon (Jacquier et al., 2017; Nam et al., 2017; Rebelo et al., 2016). The additional ~40 C-terminal amino acids in the NFH protein cause aggregation of NFH protein and NFs in the soma of transfected neurons via the formation of an aggresome (Jacquier et al., 2017; Rebelo et al., 2016). The aggregation of NFs in the

soma was shown to induce caspase-3 dependent apoptosis in transfected cells (Jacquier et al., 2017) and deprive the axons of NFs, resulting in a possible loss of function in the axonal compartment (Rebelo et al., 2016).

Mutations in *NEFH* cause a primarily adulthood onset (average ~30 years of age) rapidly progressive motor neuropathy termed as CMT 2CC (Pipis et al., 2022). However, some variability is reported in the disease onset, from 4 years of age to 40 years. In contrast to other axonal CMTs and *NEFL* CMT2E, the CMT 2CC patients also display pronounced proximal muscle weakness, which indicates that the axonal degeneration is not length-dependent (Jacquier et al., 2017; Rebelo et al., 2016). However, in severe recessive cases of CMT1F some proximal muscle weakness has been reported (Abe et al., 2009; Yum et al., 2009). CMT 2CC is a relatively severe form of neuropathy since approximately half of CMT 2CC patients reported to date needed a wheelchair after ~25 years from symptom onset (Pipis et al., 2022). Six different variants, all resulting in stop-loss, have been reported to date in 30 patients (Aruta et al., 2021; Bian et al., 2018; Ikenberg et al., 2019; Jacquier et al., 2017; Nam et al., 2017; Pipis et al., 2022; Rebelo et al., 2016).

As with mutations in *NEFL* the disease mechanism of *NEFH* variants is not clear. The primary cellular consequence is the formation of an aggresome and aggregation of NF proteins in the soma, but what pathway leads to neuron death is unclear. Nevertheless, a common hallmark with dominant and recessive *NEFL* CMT is the loss NFs or their proper organization in axons.

2.3.6 Neurofilament aggregation in other disorders

Pathological NF aggregation has also been reported in multiple neurodegenerative diseases, such as Alzheimer's disease (Ishii et al., 1979), Parkinson's disease (Goldman et al., 1983), Amyotrophic lateral sclerosis (Delisle & Carpenter, 1984), and frontotemporal dementia (Perry et al., 1987). NF aggregates can contain only NFs or they can be associated with other aggregating proteins in the above-mentioned diseases, for instance with amyloid-β, α-synuclein, TDP-43 and tau (Didonna & Opal, 2019). The initial cause for NF aggregation in neurological diseases is unclear. Nevertheless, hyper-phosphorylation of IF proteins has been suggested to lead to cytoplasmic aggregates (Sihag et al., 2007). The hyper-phosphorylation of NF proteins might protect them from proteolytic enzymes, unbalanced assembly and disassembly dynamics or lead to disrupted interaction with motor proteins (Didonna & Opal, 2019). NF aggregates causing progressive neuropathy are also seen in murine overexpression models of NF proteins (Didonna & Opal, 2019; Xu et al., 1993).

NF aggregates are more directly responsible for neurodegeneration in disorders where chaperones or ubiquitin ligases responsible for NF folding and degradation are mutated. Mutations in *HSPB1* (Zhai et al., 2007), *Heat Shock Protein Family B (Small) Member 8 (HSPB8)* (Tang et al., 2005) and *TRIM2* (Ylikallio et al., 2013) cause axonal CMT, and mutations in *GAN* (Koch et al., 1977), *DDB1 And CUL4 Associated Factor 8 (DCAF8)* (Klein et al., 2014) and *BAG Cochaperone 3 (BAG3)* (Jaffer et al., 2012) cause giant axonal

neuropathy. A common hallmark of all these disorders is the imbalance between degraded NF proteins and their proper expression and folding, and subsequent aggregation in neurons.

Past the mechanisms of NF aggregation are the pathological consequences of NF aggregates. As with *NEFL*-linked CMT, there are multiple theories on the mechanisms of disease in more general NF aggregation: disrupted axonal transport of mitochondria and other organelles (Brownlees et al., 2002; Fabrizi et al., 2007), dysregulation of mitochondrial shape and localization (Gentil et al., 2012), and energy metabolism defects (Israeli et al., 2016). Surprisingly for a protein originally characterized as solely structural, majority of these proposed mechanisms are related to mitochondria and energy metabolism in neurons warranting more mitochondria focusing studies on NF loss or aggregation.

2.4 Human motor neuron modelling

2.4.1 The motor system

Movement is controlled by several interconnected neural circuits. In a general view from rostral to caudal, the neuroanatomical regions controlling movement are the motor cortex, basal ganglia, cerebellum, brain stem and the spinal cord. For simplicity, if we bypass the cortical control in decision making, non-reflexive skeletal muscle control can be initiated by the upper motor neurons in the motor cortex, which signal to spinal motor neuron circuits in the spinal cord, which contains lower motor neurons and interneurons. Sensory feedback from muscle to the spinal cord, basal ganglia and cerebellum orchestrate coordinated muscle control with the motor cortex signal. From the spinal motor neurons, electrophysiological signal is conveyed to the muscle fibers by contact sites of axon and muscle, the neuromuscular junctions. Each muscle fiber is innervated by one motor neuron axon, but a motor neuron can innervate multiple muscle fibers. Motor neurons form motor pools in the spinal cord that control movement of a single muscle. Muscle contraction is controlled by spinal alpha motor neurons (α-MN), and spinal gamma motor neurons (γ-MN) control intrafusal muscles and muscle tone. Spinal beta motor neuron (β-MN) function is still elusive. In reflexive movement the cortical circuits are bypassed, and a sensory stimulus will activate spinal motor neurons to initiate a fast muscular response (Stifani, 2014).

In peripheral neuropathy the spinal motor neuron axons degenerate, decreasing the amount of innervated muscle fibers. This leads to the loss of muscle tone and contractility, which in turn leads to muscle atrophy. Muscle atrophy in the most peripheral parts of the body are hallmarks of axonal neuropathy (Klein, 2020).

2.4.2 Motor neuron development and specification

All tissues in the body are formed from the three germ layers, ectoderm, mesoderm, and endoderm that develop in the fertilized egg during gastrulation. All neurons are derived from the ectoderm. The ectoderm also gives rise to the neural crest and surface ectoderm derivates such as skin. Motor neuron differentiation is initiated by the specification of

neuroectoderm and formation of the neural plate from the primitive ectoderm (neurulation). Neurulation is achieved by inhibiting transforming growth factor β (TGFβ) signals through bone morphogenic protein (BMP) antagonists and activin signaling in the developing embryo. Neurulation is enhanced by activating Wnt and fibroblast growth factor (FGF) signaling. The Wnt and FGF signals induce β-catenin expression. Nuclear β-catenin inhibits BMP expression, and TGFβ antagonists such as chordin and nogging block BMP signaling resulting in the expression of neural genes, such as *SRY-box transcription factor 2* (*SOX2*) and *SRY-box transcription factor 3* (*SOX3*) (Davis-Dusenbery et al., 2014; Stern, 2005).

2.4.2.1 Spinal motor neuron progenitors

Neural tube is formed by the neuroectoderm cells by the folding of the neural plate. The neural plate and developing neural tube elongates in the rostral-caudal axis. Secretion of retinoic acid from caudal paraxial mesoderm drives the development of the spinal cord in the caudal region of the developing neural tube (Figure 7A). The neural tube is flanked by the floor plate, roof plate and the notochord, which secrete differentiation factors. Sonic hedgehog protein (Shh) is secreted from the floor plate and notochord while BMP antagonists originate the roof plate. The factors drive motor neuron specification in the dorsal-ventral axis (Stifani, 2014). Shh from the floor plate and notochord prevents the degradation of GLI-Kruppel family (GLI) proteins, which activate motor neuron differentiation signaling leading to *NK2 homeobox 2* and *9* (*NKX2.2/2.9*), *NK6 homeobox 1* and *2* (*NKX6.1/6.2*), and *oligodendrocyte transcription factor 2* (*OLIG2*) expression. Additionally, Wnt and BMP agonists secreted from the roof plate induce GLI repressor transcription leading to *paired box 3/6/7* (*PAX3/6/7*), *developing brain homeobox 1* and *2* (*DBX1/2*), and *Iroquois related homeobox 3* (*IRX3*) expression (Shirasaki & Pfaff, 2002). The combined gradient effect of GLI activation and repression leads to the initial patterning of spinal motor neuron progenitors in the ventral section of the developing spinal cord (Stifani, 2014) (Figure 7B).

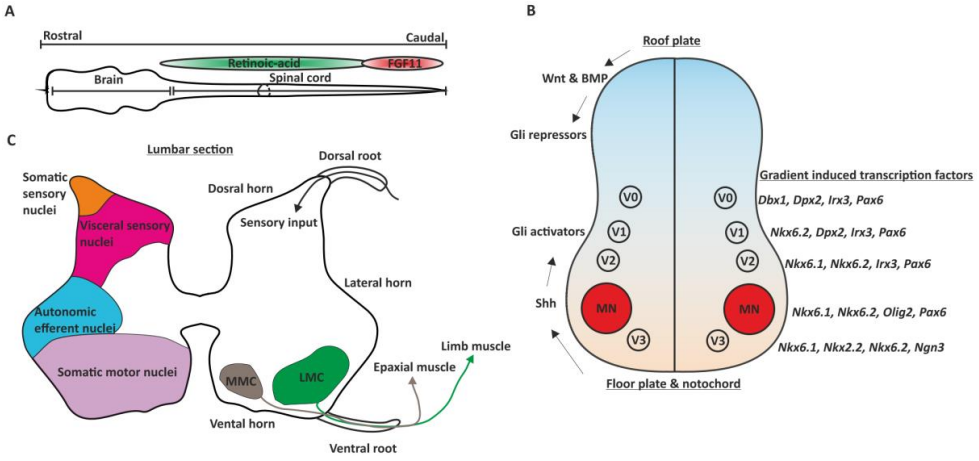

Figure 7. Motor neuron development. A. Neuron progenitors are guided towards a spinal identity in the rostral-caudal axis via gradients of retinoic acid (RA) and fibroblast growth factor 11 (FGF11). B. Gradients of Wnt activators, bone morphogenic protein (BMP) antagonists and sonic hedgehog protein (Shh) guide the differentiation of spinal neuron progenitors into motor neurons (MN) and interneurons (V0-3). Wnt and BMBs are secreted by the roof plate and Shh by the floor plate and notochord. Wnt and BMPs lead to GLI repression and Shh to GLI activation. The gradients induce the expression of cell type specific transcription factors. C. Lumbar section of mature human spinal cord. Motor neurons reside in the ventral horn of the spinal cord. Limb muscle innervating lateral motor column (LMC) motor neurons reside more distally and epaxial muscle innervating medial motor column (MMC) motor neurons localize more centrally. The left side of the section represents an estimate of different spinal neuron nuclei localization in the spinal cord. Based on Stifani et al., 2014.

2.4.2.2 Spinal motor neuron progenitor specification

Spinal neuron progenitors are further specified during development in the spinal cord, between the roof plate and the floor plate, by gradients of the above-mentioned differentiation factors to form five progenitor domains p0, p1, p2, pMN and p3. Interneurons, V0 to V3, develop in P0-P3 domains and motor neurons in pMN (Davis-Dusenbery et al., 2014; Stifani, 2014). Motor neuron progenitors expressing *OLIG2* and MN differentiation factors such as *neurogenin 2* (*NEUROG2*), acquire a post-mitotic motor neuron fate through continued retinoic acid regulation (Mizuguchi et al., 2001). This leads to *motor neuron and pancreas homeobox 1* (*MNX1* or *HB9*) expression, which stimulates its own expression and the commitment into a post-mitotic motor neuron fate (Lee et al., 2009). Furthermore, Shh secretion induces the expression of *LIM homeobox 3* (*LHX3*) and *ISL1 transcription LIM homeodomain* (*ISL1*) in motor neuron progenitors, which in turn initiate cholinergic neurotransmission gene expression (Cho et al., 2014; Ericson et al., 1992). During the development from progenitor to post-mitotic motor neurons, the cells populate the ventral horn of the spinal cord (Stifani, 2014).

2.4.2.4 Spinal motor neuron columns

In later stages of development, the motor neuron populations are distributed into columns and specific muscle innervating pools in the spinal cord (Figure 7C). The spinal cord is categorized into sections from head to tailbone, cranial, thoracic, lumbar, and sacral (Davis-Dusenbery et al., 2014). Expression pattern of different Hox-genes, determined by gradients of FGF, the growth differentiation factor 11 (GDF11), TGFβ family proteins and retinoic acid, determines the location specific identity of the motor neuron in this columnar rostral-caudal axis (Liu et al., 2001). Motor neurons innervating limb musculature, relevant in this thesis, reside in the lateral motor column (LMC) in the brachial (cervical mid-section to thoracic start) and lumbar sections of the spinal cord. These neurons have high expression of *ISL2, LHX1, HB9* and *FOXP1* (Dasen & Jessell, 2009).

2.4.2.5 Muscle targeting

Hox genes also drive the intrinsic mechanisms of MN pool specification and muscle targeting with *NKX6.1* (Dasen et al., 2005; Garcia & Jessell, 2008). Initial exit of a motor neuron axon from the spinal cord is driven by intrinsic LHX3/4 signals (Bravo-Ambrosio & Kaprielian, 2011), but afterward extrinsic signals also guide axonal targeting. How the targeting of LMC MNs is determined is still elusive, however, Ephrin-Eph and semaforin signaling has been shown to regulate axon targeting (Bonanomi & Pfaff, 2010; Dudanova et al., 2012). Correct muscle targeting is tightly regulated in the developing embryo, as indicated by knockout mouse studies. For example, *EPH Receptor A4* (*EphA4*) KO leads to increased projection into the ventral half of a limb and *Lhx1* KO leads to scrambled axon innervation in the limb. Near the target, muscle cells secrete attractive extrinsic signal molecules (i.e. glial derived neurotrophic factor, GDNF) to the growing axons, and established neuromuscular junctions provide positive feedback to axon growth and neuromuscular junction development (Bonanomi & Pfaff, 2010). Fine-tuned motor neuron gene expression patterns can be used to identify motor neuron type and innervated muscle type, for example α-MN have low expression of *estrogen related receptor gamma* (*Err3)* and high expression of *RNA binding fox-1 homolog 3* (*Rbfox3*, also called *NeuN)*. The motor neurons further develop into α, β and γ motor neurons that target muscle according to their columnar location in the spinal cord and pool identity. Alpha-MN target extrafusal muscles, γ-MN intrafusal muscles, and β-MN targeting is still elusive. In addition, the α-MN acquire characteristics of fast-firing fatigable, fast-firing fatigue-resistant and slow-firing fatigue-resistant according to their afferent muscle (Stifani, 2014).

An excess of motor neurons and axons is developed first, followed with selective pruning of connections resulting in appropriate number of motor neurons innervating target musculature. Approximately 40% of the developed motor neurons die in this pruning process (Hamburger, 1975). First, a portion of improper motor neuron "candidates" are lost during differentiation. Second, trophic extrinsic support from the target tissue provides positive feedback and survival signals to properly innervating motor neurons (Stifani, 2014).

2.4.2.6 Schwann cells

After initial muscle targeting the existing motor neuron axons grow in caliber (Leterrier et al., 2017). Schwann cells support the radial and longitudinal growth of the axons and envelope the axons for myelination or form nonmyelinated Remak fibers with the axons. The largest axons express high levels of neuregulins (NRGs), which induce the production of a thick myelin sheath by the Schwann cell (Salzer, 2015).

Compelling evidence shows peripheral axons to be metabolically supported by glia. Current studies show that in addition to myelination, Schwann cells provide trophic and metabolic support to peripheral axons (Nave, 2010; Reed et al., 2021). Myelinating cells can provide glycolytic metabolites to neurons that rely heavily on oxidative phosphorylation (Jha &

Morrison, 2018; Meyer et al., 2018). Thus, it is not surprising that axonal neuropathy can originate from neuronal, Schwann cell or combinatory origin.

2.4.2.7 Motor neuron axons

Motor neurons axons propagate the action potentials from the cell body in the spinal cord to the target muscle. Mature human motor neurons have extremely long axons, even over one meter from lumbar spinal cord to distal foot muscles in adults. Maintenance of cellular homeostasis far away from the nucleus is dependent on functional cellular transport of organelles, enzymes, and metabolites throughout the cell (Markworth et al., 2021; Schiavon et al., 2021). Furthermore, the motor axons have a specialized dynamic cytoskeleton that functions as a scaffold for axonal organelles and proteins, enables molecular transport in the axons, maintains axonal structure and protects the long processes from physical stress. The main components of the axonal cytoskeleton are microtubules, actins and NFs (Leterrier et al., 2017).

2.4.3 Human *in vitro* models of disease

Human skin-derived fibroblasts, immortalized and normal, have become one of the standard tools to study human inherited disease and molecular mechanisms. The fibroblasts can be cultured and passaged for relatively long periods providing a stable tool to investigate molecular mechanisms in human non-cancer cells. A great benefit of fibroblasts is their relatively easy derivation from any skin sample. Patients with a certain disease can be recruited and their own cells can be used to study mechanisms that are specific to their disease mutation. Similarly to cancer cell lines, fibroblasts do not provide the correct cellular context for majority of disorders. The next development in *in vitro* research came from stem cells.

2.4.3.1 Stem cells

Rodent stem cells derived from fertilized eggs set the stage for usage of human stem cells in research (Martin, 1981). It took almost 20 years from the culture of mouse embryonic stem (ES) cells to develop methods to culture the human counter parts (Thomson, 1998). Human blastocysts are used to harvest ES cells from the inner cell mass. The blastocyst stem cells are pluripotent cells, with the ability to produce cells from all three germ layers: mesoderm, endoderm, and ectoderm (Chen & Lai, 2015). With external ques researchers can guide the stem cells to differentiate into any cell type found in the body (Takahashi & Yamanaka, 2015). The stem cells, as cancer cell lines, can theoretically proliferate indefinitely in the right conditions and maintain their pluripotency. The usage of ES cells has some availability and ethical concerns since they originate from fertilized human eggs, even though they are mainly the byproduct of *in vitro* fertilization by the consent of donors.

A breakthrough in cell type specific disease modelling came in 2007 when the Yamanaka group discovered exogenous transcription factors that could be used to reprogram differentiated somatic human cells into iPSC (Takahashi et al., 2007). The factors were termed as the Yamanaka factors: *SOX2*, *kruppel like factor 4* (*KLF4)*, *POU class 5 homeobox 1* (*POU5F1*, also called *OCT3/4)* and *MYC proto-oncogene* (*MYC*). The

reprogrammed iPSCs could then be differentiated into all germ layers similarly to the human blastocyst derived ES cells. The reprogramming factors can be introduced to the somatic cells with viral transfection, plasmids, proteins, mRNA or with activation of endogenous pluripotency loci. Nonviral delivery of reprogramming factors or activation of endogenous loci is preferred to reduce the likelihood of gene integration to the reprogrammed cell genome (Takahashi & Yamanaka, 2016).

Currently researchers can take, for example, a skin biopsy from a patient with a cell type specific disorder, culture fibroblasts, reprogram them into iPSC and then differentiate those into the desired cell type specific disease models (Figure 8). Differentiated cell type specific markers and single-cell technologies can be utilized to study only the cells of interest to discover disease and cell type specific mechanisms of disease.

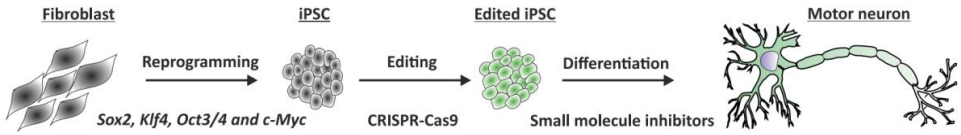

Figure 8. Motor neuron disease modelling. Patient or control fibroblasts are reprogrammed into induced pluripotent stem cells (iPSC). Patient stem cell clones are corrected, or control clones mutated with CRISPR-Cas9 to produce isogenic pairs. Gene-edited and original clones are differentiated into iPSC-MNs. The iPSC-MNs are then assayed to investigate cell type specific mechanism of disease.

Disease modelling still has multiple issues to overcome in the coming years. The differentiated cells from iPSC are more fetal-like cells than actual mature cells, so they can optimally only be used for modelling very early-onset disorders. More detailed differentiation methods must be developed to generate repeatable high purity differentiated cells. In addition, in many disorders many cell types contribute to disease. Usage of co-culture and organoid systems can overcome some of these issues.

2.4.3.2 Direct lineage conversion

A recent alternative to stem cell-based disease modelling is direct lineage conversion. Somatic cells, such as fibroblasts, can be directly reprogrammed into another somatic cell type. Exogenous transcription factors can be used to initiate differentiation programs in the already terminally differentiated cell (Vierbuchen & Wernig, 2011). For example, *achaete-scute family BHLH transcription factor 1* (*Ascl1)*, *LIM homeobox transcription factor 1 alpha* (*Lmx1a)* and *nuclear receptor subfamily 4 group a member 2 (Nurr1)* can be used to transdifferentiate mouse embryonic fibroblasts into dopaminergic neurons (Vierbuchen et al., 2010) and *neuronal differentiation 1 (NEUROD1)*, *ASCL1, POU class 3 homeobox 2 (POU3F2,* also called *BRN2), myelin transcription factor 1 like (MYT1L), LHX3, HB9, ISL1* and *neurogenin 2* (*NEUROG2)* to induce motor neuron conversion of human embryonic fibroblasts (Son et al., 2011). In contrary to iPSC reprogramming, transdifferentiation does not reset the epigenic state of the cell, therefore retaining the age of the cell. This can be an

advantage in modelling late-onset or multifactorial disorders. On the other hand, the transdifferentiated cell might retain some confounding characteristics of its original cell type through epigenetics. In addition, direct reprogramming is less efficient and reproducible than using an iPSC for differentiation (Vierbuchen & Wernig, 2011).

2.4.3.3 Stem cell editing

Another instrumental breakthrough in disease modelling was the development of clustered regularly interspaced short palindromic repeats (CRISPR)-Cas9 technology to accurately cut and edit the genome of any cell in culture (Jinek et al., 2012). Initial studies on gene engineering with phage based homologous recombination revealed the editing to be enhanced with controlled double strand breaks (DSB) near the edited site (Rouet et al., 1994). Soon after, nuclease fused zinc finger (ZFN) proteins were utilized to target DNA cleavage. ZFNs contain DNA recognition domains that can be modified to target a 9bp or 18bp specific nucleotide sequences (Kim et al., 1996). The modification of ZFN to accurately bind to specific DNA sequences is difficult and unreliable (Torres-Ruiz & Rodriguez-Perales, 2017). The development of ZNF was followed by transcription activator-like effector nucleases (TALENs) (Moscou & Bogdanove, 2009). TALENs bind DNA via a tandem repeat of 34 amino acids, with each repeat recognizing one nucleotide. The repeat amino acid sequence can be modified to recognize a specific DNA sequence. As with ZNFs, the TALENs are laboursome to generate and lack specificity, since both technologies depend on protein sequences binding to DNA sequences (Torres-Ruiz & Rodriguez-Perales, 2017). In contrast to protein-DNA interaction, CRISPR-Cas9 technology utilizes DNA-DNA Watson-Crick base-pairing with incomparably accurate specificity.

CRISPR-Cas9 is based on a bacterial immune response against bacteriophages (Barrangou et al., 2007). The genome of the bacterium *Streptococcus thermophilus* contains arrays with a conserved repeat fragment and short (20nt) recognition sequence complimentary to bacteriophage DNA, short CRISPR RNAs (crRNA). In genome editing the recognition sequence is termed as a guide. The bacteria also produce a Cas protein, which can cleave DNA. The Cas protein binds the crRNA, and the recognition sequence in the crRNA directs the Cas protein to the viral DNA to cut it. Additionally, a 20nt sequence recognized in the targeted DNA must be followed by a protospacer-adjacent motif (PAM) sequence, i.e. NGG for Cas9, for the Cas protein directed cleavage. Multiple Cas proteins have been recognized in different bacteria, but the most widely used protein in gene-editing is the Cas9 (Gasiunas et al., 2012). Using guide sequences complimentary to the genomic location of interest, researchers can use the CRISPR-Cas9 system to generate specific DNA cuts to, for example, enhance homologous recombination at the site.

The targeted DSB generated in DNA is repaired by two conserved mechanisms, homology-directed repair (HDR) and non-homologous end joining (NHEJ). In HDR a template,

exogenous homologous oligonucleotide or allele pair, is used to repair the site with a DSB. In gene-editing a single- (ss) or double stranded (ds) oligonucleotide (ODN) is introduced as a template to generate a specific chance in the genomic loci with the DSB. In contrary, NHEJ does not depend on a template to repair DSBs. It is an error prone system that integrates indels at the site, possibly resulting in frameshift variants. In gene-editing the NHEJ can be utilized to produce null-alleles, KOs (Torres-Ruiz & Rodriguez-Perales, 2017).

Recently modified nickase or dead CRISPR-Cas9 (dCRISPR-Cas9) systems have been developed that target DNA with the specific guides but do not create DSBs. Direct modification of cytosine to thymine and adenine to guanine can be achieved with base editors. Base editors are a dCRISPR-Cas9 complex guiding a nucleotide deaminase to a specific genomic location with guide accuracy. The deaminase will enzymatically convert nucleotides in the DNA in a small active site, cytosine to uracil and adenine to ionosine. Uracil is converted to thymine and ionosine into guanine upon DNA replication or repair (Kantor et al., 2020). The most recent development is the prime editing system, which utilizes a modified prime-editing guide and a reverse transcriptase bound to a CRISPR-Cas9 nickase. The prime-editing guide, which contains a targeting sequence and a new template strand with desired modifications, can then guide the reverse transcriptase to the desired location to modify the DNA in a programmable fashion (Anzalone et al., 2019). Currently, the CRISPR-Cas9 system can be used in a myriad of ways from gene activation to fluorescent labeling of DNA (Torres-Ruiz & Rodriguez-Perales, 2017).

CRISPR-Cas9 technology has made gene-editing accurate and easily accessible, however the guide directed targeting is not perfect. Off-target DSB can be generated in the genome with unpredictable consequences. Therefore, guides should be designed to minimize the amount of potential off-targets.

2.4.3.4 Motor neuron differentiation in vitro

Stem cells can be differentiated into motor neurons (Figure 9) to study the molecular mechanisms of motor neuron degeneration and axonal neuropathies in the correct cellular context. Extensive studies on murine, avian (Ericson et al., 1992) and amphibian embryonal development have revealed developmental cues of motor neurons (Jessell, 2000) that have later been mimicked to differentiate motor neurons from human stem cells. In 2009 Chambers (Chambers et al., 2009) discovered the inhibition of the SMAD signaling with two inhibitors (termed dual-SMADi), by inhibition of BMP pathway with Nogging, and TGFβ pathway with SB431542, to be instrumental in neural conversion of human ES cells. Continued development of stem cell differentiation towards a motor neuron fate revealed that early dosage of inhibitors (i.e. CHIR) inducing Wnt agonist expression increases spinal neuron conversion during dual-SMADi (Maury et al., 2015). Neuron progenitor caudalization and ventralization towards spinal motor neurons needs a balanced time dependent concentration of Smoothened agonist (SAG), which activates sonic hedgehog

signaling, and retinoic acid, as in the *in vivo* spinal cord development, for proper motor neuron differentiation (X. J. Li et al., 2005; Sances et al., 2016). The developing motor neurons depend on neuronal growth factors for survival and maturation, therefore brain derived neurotrophic factor (BDNF), ciliary neurotrophic factor (CNTF) and GDNF are typically added to differentiating motor neurons (Sances et al., 2016). Motor neuron maturation and cell cycle exit is in turn induced with inhibition of Notch signaling by gamma-secretase inhibitors (Maury et al., 2015; Sances et al., 2016).

In addition to small molecule inhibitors and growth factors, the cellular microenvironment and extracellular matrix (ECM) affect differentiation *in vitro.* Most refined protocols use suspension-based spheroid cultures in the first steps of differentiation to facilitate proper cellular connections during differentiation (Fujimori et al., 2018; Guo et al., 2017). After suspension culture the spheres are typically dissociated to achieve a homozygous 2D culture containing a network of motor neurons. Here, the motor neurons need the proper ECM for maturation and axon outgrowth, laminin. In some protocols Matrigel is used, it is an ECM protein mix containing laminin (Fujimori et al., 2018; Guo et al., 2017; Shimojo et al., 2015). In an alternative method, motor neurons can be generated by viral transfection of key motor neuron differentiation transcription factors, such as *ISL1*, *NGN* and *LHX3*, or by generating gene-edited inducible cell lines (Fernandopulle et al., 2018; Shi et al., 2018).

iPSC derived motor neurons have been successfully used to model neuropathies by multiple different labs (Table 8). NFL p.N98S variant was shown to aggregate and cause reduction in mitochondrial displacement in patient iPSC derived motor neurons (iPSC-MN) (Saporta et al., 2015). In an isogenic pair of CMTX6 iPSC-MN investigating a p.R158H variant in the *pyruvate dehydrogenase kinase 3 (PDK3*) gene revealed that the mutant protein hyperphosphorylated the mitochondrial pyruvate dehydrogenase complex (PDC). This led to mitochondrial morphology and respiratory changes, which were reversed by inhibiting PDK with dichloroacetic acid treatment (Perez-Siles et al., 2020). In a recent endeavor to determine common mechanisms of CMT Van Lent (Van Lent et al., 2021) discovered that iPSC-MN with *MFN2* p.R94Q, *NEFL* p.P8R, *HSPB8* p.K141N, or *HSPB1* p.G84R or p.P182L variants had reduced mitochondrial speed.

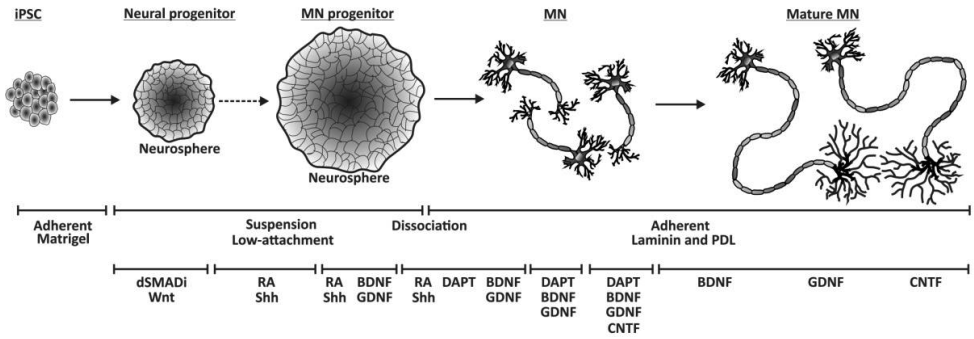

Figure 9. Motor neuron differentiation *in vitro*. Example protocol for iPSC-MN differentiation *in vitro*. Neuron progenitor development is initiated with dual-SMAD inhibition and Wnt activation in suspension culture. Retinoic acid and Smoothened agonist induce motor neuron specification. Growth factors BDNF, GDNF and CNTF support motor neuron maturation and Notch inhibition enhances cell cycle exit. Dissociation of spheres enables systematic culture conditions and neurite growth tracing. Laminin enhances neuron attachment and neurite outgrowth after dissociation. Motor neurons mature and grow elaborate neurite networks in long-term adherent culture. Based on Guo et al., 2017.

Table 8. iPSC-MN models of CMT and HMN

Reference	CMT	Gene and mutation	Main finding
(Saporta et al., 2015)	**CMT2E**	***NEFL* p.N98S**	**Aggregation and reduced mitochondria displacement**
(Van Lent et al., 2021)	**Multiple CMT2**	***MFN2* p.R94Q, *NEFL* p.P8R, *HSPB8* p.K141N, and *HSPB1* p.G84R and p.P182L**	**Reduced mitochondria speed**
(Haidar et al., 2019)	**CMT2F**	***HSPB1* p.P182L**	**Decreased autophagic flux**
(Perez-Siles et al., 2020)	**CMTX6**	***PDK3* p.R158H**	**Reduced mitochondria size**
(Sainio et al., 2018)	**CMT1F**	***NEFL* p.R367X**	**Complete loss of NFL protein**
(Kim et al., 2016)	**CMT2F**	***HSPB1* p.S135F and p.P182L**	**Reduced amount of moving mitochondria rescued with HDAC6 inhibitors**
(Manganelli et al., 2019)	**CMT2DD**	***ATP1A1* p.Pro600Ala**	**Differentiation defect**
(Maciel et al., 2020)	**Multiple CMT**	***NEFL* p.N98S and p.E396K, MFN2 p.R364W**	**Aggregation of N98S in axons**
(Alderson et al., 2021)	**CMT2F**	***HSPB1* p.P182L**	**Impaired chaperone activity and aggregation**
(Feliciano et al., 2021)	**CMT2E**	***NEFL* p.N98S**	**Aggregation can be rescued by deleting the mutant allele**
(Miressi et al., 2021)	**CMT4A**	***GDAP1* p.Ser194X**	**Reduced viability in differentiation and altered mitochondrial morphology**
(Juneja et al., 2018)	**Multiple CMT2**	***HSPB8* p.K141N, *RAB7* p.V162M, *NEFL* p.P8R, *GDAP1* p.H123R, *HSPB1* p.R127W, *MFN2* p.R94Q**	**PFN2 reduction in NEFL P8R and MFN2 R94Q**
(Perez-Siles et al., 2020)	**dHMN**	***ATP7A* p.T994I**	**Reduced basal respiration and axonal mitochondria amount**
(Shi et al., 2018) (Schwann cells and MNs)	**CMT1A**	***PMP22* gene duplication**	**Schwann cells fail to myelinate motor neuron axons**
(Mou et al., 2021)	**CMT2A2B***	***MFN2* knockdown**	**Reduced mitochondrial transport**
(Ohara et al., 2017)	**CMT2A2A/B**	***MFN2* p.H128Y and p.R94Q**	**Smaller and reduced number of neurite mitochondria**

*Haploinsufficiency in *MFN2* has not been reported to cause CMT.

Since no differentiation protocol is perfect, many protocols utilize proliferating cell toxins, mainly cancer drugs, to eliminate all non-post-mitotic cells (Perez-Siles et al., 2020). Also, some protocols use fluorescence activated cell sorting (FACS) to purify the population of

cells (Saporta et al., 2015). Transgenic tools are also used to select cells expressing marker genes under motor neuron promoters, such as *HB9* (Klim et al., 2019).

Unfortunately, there is ambiguity in the final characterization of mature motor neurons and there is yet no gold standard for motor neuron differentiation. These issues lead to variability and difficulties in reproducing data (Juneja et al., 2019). Many studies use a mix of quantitative PCR, western blot and immunocytochemistry to test markers highly expressed in motor neurons. *ISL1*, *choline O-acetyltransferase* (*CHAT)* and non-phosphorylated NFH are the commonly used markers together with *HB9*. Some spinal interneurons also express *ISL1* (Cho et al., 2014), *CHAT* is expressed by all cholinergic neurons, such as motor neurons and preganglionic autonomic neurons (Jordan et al., 2014), and non-phosphorylated NFH is high but not restricted to MN (Tsang et al., 2000). Therefore, a combinatory use of multiple markers and techniques is recommended to assay motor neuron identity. A gold standard for differentiation and validation would greatly benefit the field since it could provide more repeatable results.

3 AIMS OF STUDY

The aims of the study were divided into two main goals, first to evaluate the effectiveness of WES in the molecular diagnosis of adult patients with rare hereditary neuromuscular disease and to expand the genetic landscape of neuromuscular disease in Finland. Secondly the goal was to model the pathological mechanisms of axonal degeneration with iPSC derived motor neuron models.

The specific aims were:

- Describe new variants causing neuromuscular disorders in Finland
- Evaluate the feasibility of WES as a first-line diagnostic tool
- Evaluate the cost-effectiveness of WES
- Expand the phenotypic spectrum of *PYROXD1* myopathy
- Model axonal neuropathy with iPSC derived motor neurons
- Unravel pathological mechanisms in NFL loss

4 MATERIALS AND METHODS

The main methods are briefly described here. Detailed methods can be found from the original articles and their supplemental material.

4.1 Genetics

4.1.1 Sanger (I, II, III)

The sequenced genomic region was amplified with gene specific primers, purified and sequenced in Finnish institute for molecular medicine (FIMM) with DigDye chemistry.

4.1.2 Gene panel sequencing (III)

Next-generation gene panel sequencing for CMT diseases was performed with the HaloPlex Target Enrichment Kit (Agilent Technologies). Detailed protocol can be found in Ylikallio et al., 2014.

4.1.3 Whole exome sequencing (I, II)

Whole exome sequencing (WES) was performed at FIMM with the Illumina HiSeq2500 system in Rapid mode using HiSeq Rapid v2 kits (Illumina) as described in II. Reads were then aligned to the GRCh37 reference genome and variants called according to standard protocols.

4.1.4 Variant filtering (I, II)

Candidate variants in known pathogenic genes (OMIM or Clinvar) were filtered from WES data and assessed with ACMG standards as pathogenic, likely pathogenic, variant of unknown significance and benign. The initial filtering included variant frequency below 1.0E-03 in gnomADv2.1 variant database and a predicted consequence in the protein sequence.

4.2 SDS-page (III, IV)

Proteins were extracted from cells with RIPA-lysis buffer (Thermo Fisher) and centrifugation. Lysates were treated with Laemmli-buffer (Thermo Fisher) containing β-mercaptoethanol (Bio-Rad) and boiled. Samples were loaded into pre-cast 8-14% TGX-acrylamide gels (Bio-Rad) and run with standard settings. Proteins were transferred to 0.2 µm Nitrocellulose- membranes (Bio-Rad) with the TransBlot Turbo (Bio-Rad) transfer system. Membranes were blocked with 5%-non-fat dry milk (Valio) Tris-base-solution-with Tween20 0.1% (TBS-T). Primary antibodies were incubated over night at 8°C and secondary antibodies for 1-2h at room temperature. Antibodies used can be found in the original publications. All antibodies were diluted into 1-3% bovine serum albumin (BSA)

(BioWest). Horseradish peroxidase signal was produced with an ECL reagent (Western Lightning, Perkin Elmer) and detected with Chemidoc XRS+ (Bio-Rad).

4.3 Immunocytochemistry (III, IV)

iPSC-MN and iPSC on cover glasses were fixed with 4% paraformaldehyde (PFA) and permeabilized with 0.2% Triton x100 (Fisher). All solutions were made to phosphate buffer saline (PBS). The samples were blocked with protease-free 5% BSA (Jackson ImmunoResearch) in room temperature for 2 h. Primary antibodies were incubated overnight at 4°C and secondary antibodies 2 h room temperature. Glasses were washed with PBS Tween20 0.1% after each step. All antibodies were diluted into blocking buffer. The cover glasses were flipped on to drops of Vectashield DAPI (Vectorlabs) on microscopic slides.

4.4 Transmission electron microscopy (III, IV)

iPSC-MN on cover glasses were fixed with 2% glutaraldehyde in 0.1 M NaCac buffer and washed with the same buffer. Samples were prepared according to standard protocols by the Electron Microscopy Unit of the Institute of Biotechnology, University of Helsinki.

4.5 Imaging (III, IV)

4.5.1 Immunocytochemistry (III, IV)

Antibody-stained samples were imaged with Axio Observer Z1 (Zeiss) fluorescent microscope with 20x and 63x objectives or the Andor Dragonfly spinning disk confocal (Oxford instruments) microscope with 60x objective.

4.5.2 Live-cell (IV)

Axonal videos for organelle tracking were imaged with the Andor Dragonfly (Oxford instruments) spinning disk confocal; pinhole 40, 1 min capture, 1 frame/sec, 40x.

4.5.3 Transmission electron microscopy (III, IV)

All ultrathin samples were imaged with Jeol 1400 (Jeol) transmission electron microscope at 80 000 V.

4.6 RNA

4.6.1 Single-cell RNA sequencing (III)

Single-cell sequencing was performed for patient and control iPSC-MNs. The cells were dissociated from one patient and one control well on day 35 of differentiation using the 10x Genomics platform (10x Genomics). Detailed description can be found from Sainio et al., 2018. 349 patient iPSC derived and 96 control iPSC derived good quality iPSC-MN lineage cells were captured and used for further analysis.

4.6.2 Quantitative RT-PCR (III, IV)

The NucleoSpin RNA kit (Macherey-Nagel) was used for RNA extraction from cultured cells. Maxima first strand cDNA synthesis kit (Thermo Fisher) was used for cDNA reverse transcription. Gene specific primers were used to assay transcript levels by qPCR amplification in CFX Real-time system C1000Touch (Bio-Rad) with SYBR-green Flash (Thermo Fisher).

4.6.3 In-situ (unpublished)

RNA in-situ hybridization was performed on fixed and permeabilized iPSC-MNs following RNAscope Multiplex Fluorescent Reagent Kit Version 2 (#323100, Advanced Cell Diagnostics) instructions. The iPSC-MNs were dehydrated and rehydrated in increasing and decreasing concentrations of ethanol (50%, 70% and absolute ethanol). Probes for *NEFL* (Cat No.468671, ACDbio) with fluorophore CY3 and *TUBB3* (Cat No. 318901-C2, ACDbio) with fluorophore CY5 were used to assay *NEFL* transcript reduction and localization.

The cell nuclei were stained with DAPI and samples mounted with ProLong Gold Antifade Mountant (P36930, Invitrogen). The samples were scanned 3DHISTECH Pannoramic 250 FLASH II digital slide scanner using 40x objective and extended focus option with 5 focus levels. Andor Dragonfly spinning disc confocal imaging was used for high quality imaging.

4.7 Microfluidics

4.7.1 Organelle tracking (IV)

All microfluidic devices were coated with poly-D-lysine 50 µg/ml (Merck Millipore) and laminin 40 µg/ml (Sigma-Aldrich) before neuron plating.

Xona PDMS (Xona microfluidics) devices on cover glasses were used in organelle tracking. Day 30 iPSC-MNs, 150 000 cells plated on day 10, were incubated with Mitotracker Green FM 50 nM (Thermo Fisher) and Lysotracker Red DND-99 50 nM (Thermo Fisher) for 1h. Media was changed, and videos captured. Movies were analyzed as videos (all moving

organelles per video) and as kymographs from individual axons. Detailed data analysis protocol is available in publication IV.

4.7.2 Axotomy (IV)

For axotomy preassembled Xona 450 µm microfluidic devices (Xona microfluidics) were used. To analyze axon growth and regrowth during days 15 to 24 of differentiation the axonal compartment was imaged daily with a phase-contrast microscope (INFO). Axons were washed out by forceful aspiration in the axonal wells on day 22 to assess axonal regrowth capacity. Fiji was used to measure the area covered by axons and number of axonal crossings in the axonal compartment. Detailed data analysis protocol is available in publication IV.

4.8 Electrophysiology

4.8.1 Patch-clamp recordings (IV)

Whole-cell patch-clamp recordings were performed as in (Harjuhaahto et al., 2020) on iPSC-MNs after day 35 of differentiation.

4.9 In vitro treatments

4.9.1 Cycloheximide (III)

Cycloheximide (Merck) (200 µg/ml) was used to inhibit protein translation and nonsense mediated mRNA decay. iPSC-MNs were treated for 18h before mRNA extraction and RT-qPCR analysis.

4.9.2 Nonsense mediated decay inhibitors (IV)

Gentamicin 50-500 µg/ml (Gibco), PTC-124 25-200 µM (Merck Calbiochem) and amlexanox 25-250 µM (Merck) were used to induce translational readthrough during iPSC-MNs maturation. The cell treatment was initiated between day 14 and 32 of differentiation and continued for 48h to 12days. Neuron health was monitored during the treatments with phase-contrast microscopy. At the end of treatment, proteins were extracted with RIPA and SDS-page performed as previously.

4.10 Cell culture

4.10.1 Induced pluripotent stem cells (III, IV)

Induced pluripotent stem cells were reprogrammed from patient and control fibroblasts at the Biomedicum Stem Cell Center (University of Helsinki, Finland) by overexpression of

OCT4, SOX2, KLF4 and *MYC*. Expression of pluripotency genes was confirmed by RT-qPCR and immunocytochemistry. The iPSC were grown in feeder-free conditions with Matrigel (Corning) and E8 media supplemented with E8 supplement (both from Gibco). At ~80% confluency, the cells were passaged with 0.5 mM EDTA.

Initial iPSC karyotype was determined by the Biomedicum Stem cell center. After CRISPR-Cas9 or long time in culture the karyotype of the iPSCs was determined again. Karyotype was determined on fixed cells arrested in metaphase with KaryoMax Colecemid (Thermo Fisher) according to common practice.

4.10.2 Mixed neuronal differentiation (V)

Differentiation into a mixed neuronal population was performed as in Hämäläinen et al., 2013. Briefly, iPSCs were dissociated into low-attachment plates and cultured as neurospheres for 6 weeks in DMEM/F12 with glutamax (Gibco) and Neurobasal (Gibco) supplemented with N2 (Gibco), B27 (Gibco) and bFGF2 (Sigma-Aldrich). The spheres were then plated onto laminin coated wells and BDNF was added and FGF2 removed from the media. After 2 weeks the cells were used for experiments.

4.10.3 Adherent differentiation (III)

Adherent iPSC-MN differentiation was performed as in Du et al., 2016. A base media of DMEM/F12 with glutamax (Gibco) and Neurobasal (Gibco) supplemented with N2 (Gibco), B27 (Gibco), ascorbic acid (Santa Cruz) and penicillin/streptomycin (Lonza) was used in all stages of differentiation. Briefly, iPSC passaged onto Matrigel the following day were treated with dualSMAD inhibitors DMH1 (Merck) and SB531524 (Merck), and a Wnt/β-catenin activator CHIR99021 (Merck) for 6 days. The cells were passaged onto fresh Matrigel-coated plates and treated with lower concentrations of the above-mentioned inhibitors in addition to retinoic acid (Fisher) and Shh agonist purmorphamine (Stemgent) for another 6 days. The iPSC-MN progenitors were then dissociated (Accutase, Innovative Cell Technologies) and transferred into low-attachment plates (Corning) for a 6-day suspension phase. The progenitor spheroids were treated with retinoic acid and purmorphamine during the suspension. For spinal iPSC-MN maturation, the spheroids were dissociated with Accutase and plated onto poly-D-lysine (PDL) and laminin coated plates. In the maturation stage the media was supplemented with retinoic acid, purmorphamine, a Notch-inhibitor Compound E (Labnet) and recombinant growth factors BDNF (Sigma-Aldrich), IGF-1 (Calbiochem) and CNTF (Thermo Scientific). After 35 days of differentiation (from iPSC), the neurons were used for experiments. After spheroid dissociation only half of the media was changed during media change every 2-3days.

4.10.4 Suspension differentiation (IV)

For the suspension-based differentiation a protocol by Guo et al., 2017 was followed. The same base media was used, where penicillin/streptomycin was exchanged with Primocin (Invivogen). Briefly, iPSCs were dissociated into small clusters with EDTA and plated into low-attachment wells. For the first 2 days the base media was supplemented with dualSMAD inhibitors SB431542 (Merck) and LDN-193189, a Wnt/β-catenin activator CHIR99021 and Rho-kinase inhibitor Y-27632 (ROCKi, Selleckchem). During the next 5 days the cells were treated with retinoic acid and a Shh signaling agonist smoothened agonist (SAG, Calbiochem). On day 7 the above media was supplemented with recombinant growth factors BDNF and GDNF (Peprotech) and on day 9 a Notch-inhibitor DAPT (Calbiochem) was added to the mix. The iPSC-MN progenitor spheroids were dissociated with Accumax (Invitrogen) onto PDL and laminin coated plates on day 10 of differentiation. The plated neuron progenitors were cultured in day 9 media supplemented with ROCKi for one day after changing back to basic day 9 media until day 14 of differentiation. On day 14 the media was supplemented with BDNF, GDNF and DAPT and on day 16 growth factor CNTF (Peprotech) was added to the mix. From day 18 onwards the iPSC-MNs were cultured with base media containing the growth factors BDNF, GDNF and CNTF. After spheroid dissociation only half of the media was changed during media change every 2-3 days. The iPSC-MNs were used for experiments after 30 days of differentiation.

4.11 Genome-editing (IV)

CRISPR-Cas9 technology was used to knockout *NEFL* from control iPSC. Two guides were designed, one targeting the 5'UTR and another targeting exon 1. Guide constructs were prepared with PCR amplification as in Balboa et al., 2015. Guide pair cutting efficiency was analyzed in HEK293 cells. The cells were transfected with the guides and CAG-Cas9-T2A-EGFP-ires-puro plasmid. The target site was then amplified, and product analyzed with agarose gel electrophoresis.

Selected guides and the Cas9 plasmid were electroporated into iPSC. Green fluorescent protein (GFP) positive cells were sorted with FACS into 96-well plate wells coated with Matrigel. ROCKi and CloneR™ (StemCell Technologies) was added to the E8 media to increase cell survival. Expandable colonies were screened by PCR to clones containing excised DNA and differentiated into iPSC-MNs to confirm the loss of NFL protein.

4.12 Cost-analysis (I)

Diagnostic routine costs for cost-analysis were gathered from patients having their first clinic visit after 2010. Patients examined before 2010 were excluded since their records were inconclusive. Costs were gathered until the end of 2018. Prices were gathered from

University hospital rates and directly from other diagnostic service providers as described in Sainio et al., 2022.

4.13 Ethical statements (I, II, III, IV)

The generation and use of iPSC in this dissertation was approved by the Coordinating Ethics Committee of the Helsinki and Uusimaa Hospital District (Nro 423/13/03/00/08) with informed consent of the donor.

The collection of patient samples, sequencing and diagnostic cost gathering was approved by the ethics committee of HUS Helsinki University Hospital.

4.14 Patient recruitment (I, II)

One hundred index patients were recruited into the Neurogenomics study for WES at adult neuromuscular outpatient clinics in Finland. The selection criteria included patient age above 16 at date of testing, no previous genetic diagnosis, and no indication to a common genetic etiology. The criteria for selection are described in detail in Sainio et al., 2022.

Two patients recruited into the Neurogenomics cohort are described in detail in article II. Two additional patients with a previously undiagnosed neuromuscular disease were included through a collaboration with the Tampere University hospital.

4.15 Single Molecule Array (IV)

NFL was measured from patient serum as in (Järvilehto et al., 2021) with the same controls.

In the cell experiments, 1 ml of media was collected from day 30 iPSC-MN. Media were prepared according to standard protocols and NFL concentration measured with the Single molecule array (Simoa) HD-1 analyzer (Quanterix, Billerica, MA) and the NF-Light Advantage Kit (ref. 103186) according to manufacturer´s instructions (Quanterix, Billerica, MA). As a positive control of NFL release from degenerating neurons the iPSC-MN were treated with Vincristine (Sigma-Aldrich).

4.16 *PYROXD1* yeast assay (II)

The *PYROXD1* p.Tyr354Cys variants oxidoreductase capacity was assayed in yeast challenged with hydrogen peroxide. Detailed description can be found in Sainio et al., 2019.

4.17 Statistics (I, II, III, IV)

Graphpad Prism (Graphpad Software version 7-9) was used for statistical analysis. Details can be found in the respective articles. One-way ANOVA with Dunnett's multiple comparisons was used in unpublished analysis reported here. A p-value of less than 0.05 was termed as significant.

5 RESULTS AND DISCUSSION

5.1 Efficiency of clinical exome sequencing in rare hereditary neuromuscular diseases of adults (I, II)

Next-generation sequencing technologies are becoming more widely used in a clinical setting, but the effectiveness of first-line clinical WES in the diagnosis of adult patients with neuromuscular diseases is yet to be established. We gathered a cohort of one hundred adult neuromuscular and neurological patients and sequenced their exomes with WES to assess the usability of WES as a first line diagnostic tool (Figure 10). Additionally, we gathered clinical costs of 60 patients to evaluate the cost-efficiency of WES. Our cohort consisted of patients with myopathy, neuropathy, ataxia, spasticity, and parkinsonism. We divided the cohort into neuronal and myopathic phenotypes for analysis. The severity, progression of disease and age of onset varied significantly in the cohort.

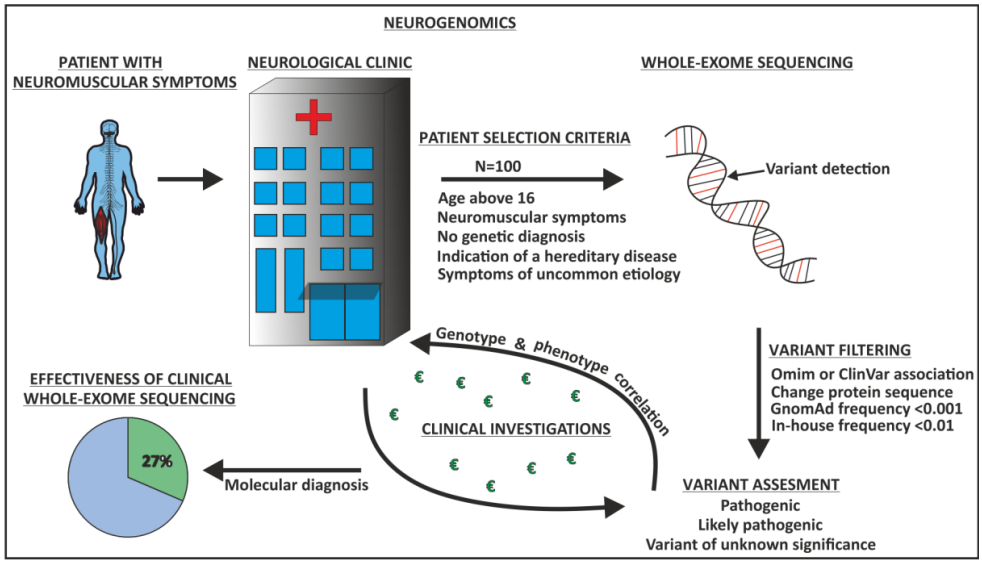

Figure 10. Clinical whole exome sequencing. Overview of the Neurogenomics project where one hundred adult patients with suspected genetic origin of neuromuscular and neurological disease were recruited and their exomes sequenced by whole exome sequencing.

5.1.1 Diagnostic yield of clinical WES (I)

The total diagnostic yield of WES was 27% in the cohort. Fourteen patients were found to carry a pathogenic variant in a total of 11 genes (*ANO5*, *CHCHD10*, *CLCN1*, *DES*, *DOK7*, *FKBP14*, *POLG*, *PYROXD1*, *SCN4A*, *TUBB3*, and *TTN*) and thirteen patients had likely pathogenic variants in 12 genes (*ABCD1*, *AFG3L2*, *ATL1*, *CACNA1A*, *COL6A1*, *DYSF*, *IRF2BPL*, *KCNA1*, *MT-ATP6*, *SAMD9L*, *SGCB*, and *TPM2*). A detailed clinical description

of all patients with new likely pathogenic variants was gathered to expand the clinical representation of given variants. Additional evidence was gathered for most patients who had the likely pathogenic variants to strengthen the genotype to phenotype correlation. For example, the muscle biopsy of the patient with *DYSF* variant was negative for DYSF staining confirming an unstable protein product, and the two patients with variants in *ABCD1* had elevated very long-chain fatty acids in their blood typically seen in X-linked adrenoleukodystrophy or adrenomyeloneuropathy. A new variant was discovered in the *MT-ATP6* gene even though the mitochondrial genome was not targeted by WES. The new variant was subsequently modelled in detail with iPSC-MN in our laboratory (Kenvin et al., 2021).

The diagnostic yield achieved in our study is comparable to previously published reports with similar patient cohorts (Eratne et al., 2021; Nagappa et al., 2018; Posey et al., 2016). For example, in a recent study containing a European cohort of 126 patients with predominantly motor peripheral neuropathy symptoms and age of onset above 35 years, 18.3% received a molecular diagnosis (Senderek et al., 2020). However, comparing WES yields between reports is difficult, since patient selection criteria influences the final yield. We chose patients with only suspected genetic origin and did not exclude patients with previous nonconclusive genetic investigations. In addition, we excluded patients who displayed clinically obvious symptoms to a specific genetic cause, such as neuropathy patients with signs of *PMP22* duplication. In other words, we chose patients with difficult to diagnose symptoms who would most benefit from the WES study. A higher diagnostic yield was achieved in a sub-cohort of patients with a positive family history (positive history 14/53 and negative history 9/73) (Senderek et al., 2020). Furthermore, our patient cohort was heterogeneous with patients with myopathy, neuropathy, ataxia, spasticity and parkinsonism. Thus, a higher diagnostic yield could be achieved with rigorous selection of patients most likely to receive a molecular diagnosis with NGS (Splinter et al., 2018) and with the selection of disease specific patient cohorts (Hartley et al., 2018; Reddy et al., 2017). However, our study was focused on the effectiveness of WES as a first-line diagnostic tool in a heterogenous cohort.

With almost one third receiving a molecular diagnosis with WES in our difficult-to-diagnose cohort, we find WES to be an efficient tool in diagnosis. WES as a first-line diagnostic tool would have shortened the diagnostic odyssey of many patients. Some of our cases received the molecular diagnosis after more than 35 years from disease onset.

We could not exclude the presence of mitochondrial variants, copy-number variations, repeat expansions, or intronic mutations in our patients, since WES cannot reliably detect these. With the usage of WGS we could detect all these variant types. In addition, there might be variants not detected in some exons by WES, since its exon coverage is not 100% (Gorcenco et al., 2020).

Since the clinical phenotypes, disease severity and onset largely vary in neuromuscular diseases, the usage of NGS should and is increasingly used in the clinic. Using NGS as a first-line diagnostic tool for all patients with suspicion of a hereditary origin would most likely save the clinician's valuable time and potentially provide a faster molecular diagnosis to patients. However, such widespread usage of NGS in the clinic would need further collaboration with geneticists and genetic counselors.

5.1.2 Determinants of diagnosis (I)

Clinical and WES data of the patient cohort was used to analyze the factors that contribute to the likelihood of finding a molecular diagnosis. We collected clinical data from all patients to evaluate the effect of disease onset and phenotype on finding a molecular diagnosis. Patients with an earlier disease onset were more likely to receive a molecular diagnosis. The median age of onset of "solved" (patients with pathogenic or likely pathogenic variants) cases was 29 years and "unsolved" (only VUS or benign findings) 40 years (p=0.034, Student's T-test). Furthermore, 33% of cases were solved in disease onset category 0-19 years, and only 13% of cases were solved in the disease onset +50 years of age category (Figure 11). Diagnostic rate was not affected by patient disease category, gender or time spent in clinic before WES.

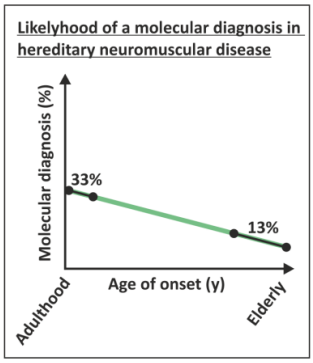

Figure 11. Molecular diagnosis of adult patients. The likelihood in discovering a genetic origin of disease decreases with increasing age of disease onset. Elderly patients can have multiple contributing variants and accumulation of acquired characteristics erroneously suspected to have a monogenic origin.

In studies with similar cohorts, a higher proportion of solved cases was also achieved for patients with an earlier onset of disease and whose parental samples were readily available (Eratne et al., 2021; Senderek et al., 2020; Trujillano et al., 2017). Furthermore, adult-onset neuromuscular disease patients can have more ambiguous symptoms, which are harder to separate from acquired characteristics or from the combinatory effect of polygenic risk factors in comparison to early-onset patients (Bis-Brewer et al., 2020). Additionally, some variants with reduced penetrance and presence in genetic databases of presumably healthy individuals can be erroneously termed as benign in databases (Richards et al., 2015). Collectively, early WES is effective in molecular diagnosis of adult-onset patients. Further consideration is needed to determine the usability and effectiveness of WES in elderly-onset cohorts of suspected genetic origin, because the diagnostic yield of NGS reduces with

disease onset. Possibly when WES costs reduce and analysis becomes more streamlined, NGS technologies will become effective in all age categories.

5.1.3 Variants of unknown significance (I)

Eighteen patients were found to carry 23 variants of unknown significance. The variants were of interest since they were in genes associated with neuromuscular diseases and their frequency in the general population was low. Their evidence of pathogeny was not high enough to categorize them as likely pathogenic, because the patient phenotypes did not completely match with what has been commonly reported for the disease gene.

As an example, a 75-year-old female with axonal sensory polyneuropathy and small fiber neuropathy had heterozygous variants in two genes: p.Val1257Phe in *SCN9A* and p.Met536Val in *ATL3*. The former was absent from gnomADv2.1, while the latter had a carrier frequency of 4.01E-6 and was marked as benign in gnomADv2.1. Variants in *SCN9A* cause dominantly inherited small fiber neuropathy and variants in *ATL3* cause dominant hereditary sensory neuropathy type IF (Faber et al., 2012; Kornak et al., 2014). Neither of the identified variants had been previously associated with disease. The patient's symptoms started at age 60 with progressive gait and balance problems. In ENMG study a severe sensory neuropathy was detected. Brain MRI showed multiple small ischemic lesions at left cerebellar cortex without signs of brain atrophy. Her younger brother had the *SCN9A* variant but not the *ATL3* variant. He has type 2 diabetes, high blood pressure, as well as normal pressure hydrocephalus that was treated with shunt operation at age 62 years. In neurography he showed signs of sensory neuropathy with absent sural nerve responses. Thus, the data collected is not sufficient to pinpoint the causative variant in the index patient, mainly because we could not determine the origin of neuropathy in the brother. Is the neuropathy caused by the *SCN9A* variant or type 2 diabetes? Future sequencing studies can reveal more patients with variants in the same genes and similar clinical representation as seen in our study. Thus, with multiple corroborating studies, VUS findings can be later described as pathogenic (Ewans et al., 2018). The opposite is also possible, the variants can be found in the population randomly with no association to a specific phenotype making them benign.

5.1.4 Genetics in Finland (I)

We found variants in three genes commonly causing neuromuscular symptoms in Finland, even though we aimed to exclude common Finnish hereditary diseases in the cohort, therefore the specifics of the analyzed population are instrumental in variant filtering, interpretation, and discovery. The Finnish population has a different genetic background as other Europeans. Finland is in a secluded northern location and its population has gone through historic genetic founder effects. A relatively small group of people settled into Finland and migration towards Finland has been limited. The reproduction of the small initial population has resulted in reduced genetic variability and enabled genetic drift in this

population. Additionally, the small initial population has gradually moved northward along Finland establishing further isolated settlements along the way. This seclusion and small populations have led to Finland having its own group of hereditary diseases, nearing 40, only seen with people of Finnish ancestry. The group of diseases, termed Finnish Disease Heritage (FDH), are caused by the enrichment of recessive variants and persistence of late-onset dominant variants (Norio, 2003a, 2003b; Salmela et al., 2008). The Finnish heritage also results in some conditions common in Europe being rarely or not at all found in Finland, such as phenylketonuria (Pastinen et al., 2001). Even though the FDH conditions are mostly specific to Finland they are still rare inside Finland.

In this study, we found Finn-major *TTN* variants in two patients with myopathic symptoms that nevertheless differed from typical tibial muscular dystrophy (TMD). The dominant Finn-major variant is a deletion and insertion of 11 nucleotides in the last exon of *TTN* (Hackman et al., 2002). In our cohort the symptoms of one of the patients were evaluated to be too severe to be solely caused by the *TTN* variant, but our analysis revealed no other contributing variants. The second patient with the Finn-major variant had a myopathic phenotype including cardiomyopathy, of which the latter is not typical for TMD. As with many FDH disorders other variants in the causative gene are found around the globe causing similar symptoms. The Finn-major variant causes a relatively mild late-onset distal muscular dystrophy, mainly affecting lower legs. Homozygosity of the variant causes a severe limb-girdle muscular dystrophy. The variant is responsible for the relatively high prevalence of TMD in Finland (Savarese et al., 2016). Haplotype analysis indicates the variant to originate from a common ancestor (Hackman et al., 2002). The *TTN* gene is large with thousands of rare benign variants and can sometimes be overlooked in analysis. In Finland special care must be taken to properly align and regard the Finn-major variant in the *TTN* gene, since the 11nt deletion and insertion can be erroneously aligned as separate variants in the 11nt region that can resemble common benign variants or errors in sequencing.

The *CHCHD10* p.Gly66Val variant was found in our cohort in a patient with a clinical representation of axonal CMT rather than spinal muscular atrophy, normally associated with p.Gly66Val (Penttilä et al., 2015). A phenotype clinically resembling CMT has also been previously described for *CHCHD10* Gly66Val (Auranen et al., 2015), but it is a rare cause of CMT even in Finland. The variant has so far only been found in people with Finnish ancestry with an estimated frequency of 12/100 000. The “healthy” carriers will most likely develop the disease at a later point in life. The Finnish founder variant in the *CHCHD10* gene is most common in the Northern Karelia region. The dominant missense variant causes late-onset spinal muscular atrophy Jokela type (Jokela et al., 2011; Penttilä et al., 2015). The disease causes slowly progressive weakness in distal and proximal limb muscles. Both the *TTN* variant and Gly66Val in *CHCHD10* cause a disease with dominant inheritance, and the symptom onset is late in adulthood. The late-onset of these dominant variants might

explain how the mutations persists in the population and why they are present in "healthy" people in variant databases.

In addition, a common Finnish *POLG* variant p.W748S causing mitochondrial recessive ataxia syndrome (MIRAS) with a carrier frequency of up to 1:125 in Finland and Norway (Hakonen et al., 2007; Van Goethem et al., 2004) was overlooked in initial clinical evaluation of a patient with ataxia, possibly because a dominant inheritance was suspected based on the patient's family history. However, the severity and onset of disease in MIRAS varies significantly making solely clinical diagnosis difficult (Hakonen et al., 2005). Even common and well characterized pathogenic variants in a population can present with varied clinical symptoms, thus the early use of unbiased NGS should be used for efficient molecular diagnosis. In variant filtering the high carrier frequency of common recessive variants and dominant late-onset disease causing variants should not be cause for variant exclusion.

Somewhat surprisingly no variants were detected in the recently discovered, and therefore previously overlooked, adult-onset neuromuscular disease genes: *MME,* causing recessively and dominantly inherited neuropathy and ataxia (Auer-Grumbach et al., 2016; Depondt et al., 2016; Higuchi et al., 2016), or *SORD*, which was discovered to be one of the most common causes of recessive hereditary neuropathy (Cortese et al., 2020).

5.1.5 Early clinical WES is cost effective (I)

In addition to clinical and genetic data, we collected the diagnostic costs from 60 cases in our cohort who entered the clinic after 2010. The diagnostic costs accumulated linearly during time in clinic without a molecular diagnosis. From the cost data we did a simple cost analysis to determine whether WES as a first-line tool would be cost-effective. In our simplistic model we speculated a molecular diagnosis to remove all other diagnostic costs and WES to replace all the other genetic costs in the cohort. In this hypothetical situation we concluded that WES could cost 2500e to be a cost-effective first-line diagnostic tool in this patient group.

In reality all costs categorized as diagnostic in our study would not be saved by a WES finding, since some diagnostic tools are additionally used to monitor the course of disease and some are used to confirm the clinical diagnosis. For example, increase in very long chain fatty acids (VLCFA) in patient serum can be used to confirm the pathogenicity of an *ATP binding cassette subfamily D member 1* (*ABCD1)* variant. On the other hand, the method used to find a pathogenic variant (WES, WGS, Sanger) does not affect the diagnostic procedures routinely done to neuromuscular disease patients. An early WES could significantly reduce the procedures done after the initial clinical investigations and therefore save costs. Studies of large cohorts of suspected hereditary disease patients have suggested the cost-effectiveness of WES (Córdoba et al., 2018; Schwarze et al., 2018; Stark

et al., 2019). We show that even in difficult-to-diagnose adult cases WES is likely to be cost-effective in the neuromuscular clinic in Finland.

As sequencing costs decline, bioinformatic tools develop and variant data aggregates, WES and WGS will be increasingly cost-effective and essential in clinical diagnosis of monogenic diseases. Researchers and clinicians should push the usage of NGS technologies in the clinic even though the initial costs seem high, since the diagnostic costs of undiagnosed patients accumulate. Faster and higher diagnosis rate would potentially be achieved by first-line sequencing of all neuromuscular patients with suspected hereditary origin.

5.2 Variants in *PYROXD1* cause late-onset LGMD (I, II)

5.2.1 Finnish patients broaden the phenotypic spectrum of *PYROXD1* variants (I, II)

Two Finnish patients with LGMD were discovered to carry variants in the *PYROXD1* gene as part of the WES cohort. The variants were found upon reanalysis of the data, following the addition of variant data reported by (O'Grady et al., 2016) to Omim (https://omim.org/) and ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/). Similar success of WES data reanalysis has been reported by others (Ewans et al., 2018). Two additional Finnish patients with late-onset LGMD and *PYROXD1* variants were identified through collaboration. The discovery of *PYROXD1* variants in late-onset patients with LGMD-like features was unexpected since variants in the gene were previously only associated with pediatric-onset myopathy with myofibrillar features (O'Grady et al., 2016). In the original report missense and insertion variants were shown to cause early-onset slowly progressive muscle weakness in a total of nine individuals from five families. The disease manifested earliest at infancy and all patients had symmetrical weakness in proximal limb musculature. All patients in the original study were still ambulant at the time of publication but had difficulties in running and raising stairs. Muscle MRI revealed fat replacement in the proximal musculature and histopathological biopsies showed mixed features of multiple myopathies (O'Grady et al., 2016).

The Finnish patients' disease onset was from 10 years of age to 49 years and symptoms progressed slowly. Therefore, the Finnish patient's disease onset was decades later and with slower progression than the ones reported earlier. Only one patient needed a wheelchair from the age of 66 years. The Finnish patients had mild weakness in distal musculature and more pronounced weakness in girdle muscles seen as reduced hip flexion strength. In addition, no myofibrillar features were seen in the Finnish patients. A consistent finding of decreased respiratory function was discovered. Magnetic resonance imaging (MRI) and electromyography (EMG) revealed atrophy and fatty replacement in lower limb muscles and myopathic changes in proximal and distal musculature. All patients had dystrophic changes in biopsies. Out of these four, three patients were homozygous with a previously described variant p.N155S and one patient compound heterozygous with p.N155S and a

novel variant p.Y354C. Segregation of the variants in the patients' families confirmed myopathy only in patients with two mutant alleles. Heterozygous carriers were unaffected.

Our and later findings in *PYROXD1* highlight the variability of clinical features even with the same variants. The current collection of patients is not sufficient for comprehensive analysis of modifiers of the severity of *PYROXD1* myopathy. However, there is an indication of a less severe and later onset disease with missense variants and an earlier onset and more severe phenotype with nonsense and splice site variants. Interestingly, no splice site variants are found in a homozygous state indicating their lethality (Daimagüler et al., 2021). Environmental effects, diet, genetic background, and mutation type could possibly explain the heterogeneity in *PYROXD1* patients. Studies of Turkish (O'Grady et al., 2016) and Sudanese of Arab ancestry (Saha et al., 2018) *PYROXD1* patients have shown a more severe disease with earlier onset compared to the Finnish patients in our study, suggesting a milder disease in patients of Northern European ancestry. However, additional reports after our publication on *PYROXD1* describe more patients with a later onset and slow disease progression in patients with Persian-Jewish descent (Woods et al., 2020) as well as patients of French origin with a neonatal disease onset (Lornage et al., 2019) and further Turkish patients with early onset (Daimagüler et al., 2021). 23 patients with *PYROXD1* variants are reported to date.

No determining conclusions can be made from variant location in the protein since the most common variant, homozygous p.N155S, is reported to have a disease onset from 2 years of age to the fifties. All six Turkish patients with homozygous p.N155S have a severe early-onset myopathy with facial weakness, speech difficulties and neuropathy (5/6), while eight non-Turkish patients with the same variants have a late-onset myopathy. The late-onset patient group only has isolated cases with facial weakness and speech difficulty (Daimagüler et al., 2021).

The most recent report (Daimagüler et al., 2021) suggests the p.N155S variant to be a Turkish founder mutation, thus the haplotype of non-Turkish patients with the variant should be investigated to confirm the variant origin.

Earlier disease onset is associated with need of ambulatory help, facial weakness and swallowing difficulties and the later onset phenotype with overall milder symptoms. The common respiratory issues of the four Finnish late-onset cases are not seen in all elderly patients, but all in all about half of the late-onset patients were reported with respiratory insufficiency. Moreover, approximately half of all patients have shown signs of neuropathy further complicating the phenotypes behind *PYROXD1* variants (Daimagüler et al., 2021). In conclusion, *PYROXD1* variants cause a spectrum of myopathic symptoms and should therefore be included in the genetic analysis of any hereditary myopathic symptoms.

5.2.2 PYROXD1 function (II)

PYROXD1 is an oxidoreductase expressed in the muscle, and is thought to regulate the redox state of other proteins. However, the exact function of the protein is still elusive.

Nevertheless, the protein is known to be a flavoprotein and a member of class 1 nuclear-cytoplasmic pyridine nucleotide-disulphide reductases (PNDR1s), which play crucial roles in metabolism and regulation of the oxidative state of proteins. PNDR1s catalyze the reducing reaction of oxidized pyrimidine residues (Argyrou & Blanchard, 2004; O'Grady et al., 2016). The reaction is dependent on bound flavin adenine dinucleotide (FAD) and oxidation of the cofactor nicotinamide adenine dinucleotide (NADH) to NAD (Argyrou & Blanchard, 2004). PYROXD1 differs significantly from the other five members of the PNDR1-family, because it contains a unique NADH-dependent nitrile reductase and a shared pyridine nucleotide-disulphide oxidoreductase domain. Furthermore, the protein does not contain a PNDR-family consensus redox active site in its oxidoreductase domain, or a C-terminal dimerization domain shared by other PNDRs (O'Grady et al., 2016).

In a recent report *PYROXD1* was shown to function in protecting a mammalian metalloenzyme (RTCB) against oxidation (Asanović et al., 2021). RTCB is the catalytic unit of the transfer RNA ligase complex (tRNA-LC). RTCB's function is regulated by the inhibitory effect of NAD(P)H. Unhindered tRNA-LC function is essential because it ligates non-canonically spliced precursor tRNAs that are needed in mRNA translation. PYROXD1 oxidates RTCB bound NAD(P)H to NADP+ and thus enables its continued function. The authors conclude that the reduction or complete loss of PYROXD1 in physiological and oxidative conditions leads to reduced ligation of noncanonically spliced RNAs (Asanović et al., 2021).

5.2.3 Mechanism of disease in *PYROXD1* myopathy (II)

PYROXD1 localizes to the nucleus in human derived *in vitro* models and to sarcomeric regions in zebrafish (O'Grady et al., 2016). PYROXD1 is highly conserved in evolution with 63% identity and 78% similarity between human and zebrafish counterparts. Morpholinos of zebrafish reducing *PYROXD1* expression resulted in swimming difficulties and muscular disruption. The experiments on human patient samples, cells and zebrafish models revealed the importance of functional PYROXD1 protein in muscle cells (O'Grady et al., 2016).

The general PYROXD1 oxidoreductase capability was assayed in yeast grown in oxidizing conditions. The endogenous yeast oxidoreductase was knocked out and complemented with human wild-type or patient mutation carrying gene. In yeast the loss of endogenous oxidoreductase and supplementation with mutated *PYROXD1* (p.N155S, p.Q372H or p.Y354C) decreased the proliferation capacity of the cells in the presence of hydrogen peroxide. Human wild-type PYROXD1 was able to rescue the loss of yeast endogenous oxidoreductase (O'Grady et al., 2016; Sainio et al., 2019).

In addition to the reduced PNDR-like oxidoreductase capacity, variants in the *PYROXD1*-gene (p.N155S and p.Q372H) reduced its ability to protect the tRNA-LC from inhibition. The variants decreased the oxidation of NAD(P)H to NADP+ dramatically. This led to the accumulation of spliced tRNAs and reduction in ligated tRNAs. Protein structure modelling

predicted the N155 amino acid to be essential for NAD(P)H positioning, thus its mutation to serine suggests that improper positioning of NAD(P)H is behind the reduced activity (Asanović et al., 2021).

Further studies in human muscle cells, either primary or iPSC-derived, should be carried out to reveal which of the PYROXD1 functions is more essential in the disease process in the affected tissue. If either phenotype is replicated in *in vitro* models, general antioxidants could be tested whether they ameliorate symptoms from the overall reduced oxidoreductase capacity, and direct use of NADP+ as an antioxidant could rescue the function of tRNA-LC in muscle. The existing morpholino zebrafish (O'Grady et al., 2016) could be used to determine the levels of ligated tRNAs to validate the findings of the most recent study (Asanović et al., 2021) in an *in vivo* system. Additionally, different tissues could be analyzed in the zebrafish to discern if the reduction in ligated tRNAs would be tissue-specific. In conclusion, the reduced oxidoreductase capacity and tRNA-LC function caused by *PYROXD1* variants present druggable pathways. A high antioxidant and NADP+ inducing diet could possibly be a low cost and tractable disease ameliorating treatment for *PYROXD1* patients, but remains to be tested. Of note, it would be interesting to measure NADP+ levels from patients with the common p.N155S variant to evaluate if the disease severity correlates with NADP+ levels and to investigate if the differences in disease severity between Turkish and Finnish patients can be explained by genetic or environmental factors.

5.3 Early-onset CMT2 caused by nonsense variants in *NEFL* (III)

A novel *NEFL* variant was discovered in a panel sequencing of neuropathy disease genes in a patient with early-onset axonal CMT. The homozygous nonsense variant in the second exon of the gene led to the introduction of a premature stop codon p.R367X. This was the fourth nonsense variant discovered in *NEFL* leading to CMT. Segregation analysis revealed both parents to be healthy heterozygous carriers of the variant. In addition to the index patient, her affected brother was also homozygous for the variant. The variant was found to have a low carrier frequency in the Finnish population (GnomAd v2.1, 0.00052).

The index patient represented a severe form of CMT resulting in the loss of ambulation at the age of 25 years. Overall, the neuropathy progression was slow but eventually led to clear muscle atrophy in the hands and feet, worsening towards the distal extremities. Nerve conduction velocities were reduced in the arms and not measurable in the legs. Furthermore, the patient had mild cognitive impairment being approximately two years behind in school. The index patient's brother was similarly affected but less severely, being ambulant at the age of 27 years. Neither of the patients had sensory symptoms.

The other reported patients with nonsense *NEFL* variants are clinically similar, summarized in Table 9. The patient with p.E140X had an early disease onset (before the age of 10) but was still ambulant with a waddling gate at the age of 44 years. He had atrophy in distal

musculature, reduced NVC and undetectable sensory action potentials. Furthermore, his brother was similarly affected and became wheelchair bound at the age of 40 years, but unfortunately the brother's DNA was not sequenced (Abe et al., 2009).

Four siblings with p.E210X variants in a consanguineous family showed consistently severe progressive neuropathy starting from infancy. All patients were evaluated in childhood and were still ambulant with aids at the time of article writing (aged from 13 to 23 years). They displayed progressive atrophy and weakness in distal musculature with reduced NCV. Additionally, some sensory symptoms were reported in these patients and one patient displayed learning difficulties (Yum et al., 2009).

A patient with p.E163X had infancy-onset progressive neuropathy similar to the p.E210X patients. At the time of examination (13 years of age) the patient had reduced NVC, distal muscle atrophy and weakness but no sensory symptoms. She was able to walk with an aid (Fu & Yuan, 2018). All patients with homozygous nonsense variants had *pes cavus* (Abe et al., 2009; Fu & Yuan, 2018; Sainio et al., 2018; Yum et al., 2009). The large myelinated axons were lost in the sural nerves of one patient with p.E210X (Yum et al., 2009) and the patient with p.E163X (Fu & Yuan, 2018). The loss of large myelinated axons can explain the reduction in NCV without the loss of myelin. In conclusion, nonsense variants in *NEFL* cause an early-onset progressive neuropathy with reduced NCV. A more clinically diverse, later onset disorder with no reduction in NCV is seen with dominant variants in *NEFL* (Stone et al., 2021). The molecular consequences of NFL loss are therefore likely more severe than those of aggregating NFL. Still both variant types lead to similar neuropathy suggesting common pathological mechanisms, such as NF loss or dysregulation in axons.

Table 9. Nonsense variants in *NEFL.*

Report	Variant	Patients	AoO	Examination	Muscle weakness	Sensory symptoms	Ambulation (age)	Median NCV(m/s)	CMAP (mV)
(Sainio et al., 2018)	**p.R367X**	**2***	**Infancy**	**Adulthood**	**Severe distal, mild proximal**	**NO**	**Wheelchair (25y)**	**21**	**-**
(Fu & Yuan, 2018)	**p.E163X**	**1**	**Infancy**	**Teenage**	**Moderate distal**	**NO**	**With ankle-foot orthoses (10y)**	**24**	**2.2**
(Abe et al., 2009)	**p.E140X**	**2***	**Childhood**	**Adulthood**	**Severe distal**	**YES**	**Waddling gait with cane (44y)**	**13.8**	**-**
(Yum et al., 2009)	**p.E210X**	**4**	**Infancy**	**Teenage**	**Severe distal, mild proximal**	**YES**	**With ankle-foot orthoses (13-23y)**	**12-25**	**1.2-2.1**

- Not measured, *affected sibling not studied in detail

5.3.1 Isogenic cells lines reduce variability (IV)

Using patient iPSC-MNs with a homozygous nonsense mutation p.R367X, we demonstrated for the first time that the mutation leads to a complete absence of NFL in neurons (Sainio et al., 2018). The patient-specific iPSC-MNs were then compared to non-isogenic controls, which are not ideal. To overcome the issues of comparing cell lines with different genetic backgrounds to each other, we generated isogenic *NEFL* KO iPSC lines with CRISPR-Cas9 by using guides targeting the 5'UTR and exon 1 of *NEFL*. Several KO iPSC lines with good growth characteristics and cut DNA were generated (Table 10). RT-qPCR revealed approximately 10% of *NEFL* mRNA expression in all *NEFL* KO cell lines even when no NFL protein was detectable in immunoblotting. Two of the analyzed cell lines were shown to be heterozygous knockouts by WB. Two full KO iPSC lines with good growth characteristics were selected for further studies.

Table 10. Screened *NEFL* patient and CRISPR-Cas9 KO iPSC-lines

Origin	iPSC line	NFL protein	*NEFL* mRNA	Use/exclusion
Patient 1	**130.1**	**Undetectable**	**~10%**	**III and IV**
Patient 1	**130.2**	**Undetectable**	**~10%**	**III**
Control 1	**IIF1**	**Undetectable**	**~10%**	**IV**
Control 1	**IIC2**	**Undetectable**	**~10%**	**IV**
Control 1	**IIE8**	**Undetectable**	**~10%**	**Backup**
Control 1	**IIIC3**	**Undetectable**	*****	**Slow iPSC growth**
Control 1	**IIID3**	**~30%**	*****	**Heterozygote KO**
Control 1	**IIIB7**	**Undetectable**	*****	**Slow iPSC growth**
Control 1	**IE9**	**Undetectable**	**~10%**	**Poor differentiation**
Control 1	**IB11**	**Undetectable**	**~10%**	**Backup**
Control 1	**IIIH7**	**~30%**	**~50%**	**Heterozygote KO**

*not measured

Pluripotency of isogenic *NEFL* KO and patient cell lines was confirmed by differentiation as well as with TRA-1-60 and Nanog immunostaining. Karyotype of cell lines used in further studies was determined. IIC2 *NEFL* KO cell line was revealed to have deletion in chromosome 18 (46, XY, del(18)(q12q22)). Since the IIC2 cell line was the only one with a karyotype abnormality, we can speculate the alteration to originate from the single cell stage after CRISPR-Cas9 and FACS. Parental control, IIF1 KO and patient cell lines had a normal karyotype.

In retrospect, using more KO cell lines and a corrected patient cell line would make disease modelling more accurate. In addition, karyotype should be determined at the cell line selection stage and preferably again during prolonged iPSC culture. However, having three genetically different cell lines all missing NFL protein makes the shared identified differences robust.

5.3.2 Complete loss of NFL protein (III, IV)

With the two differentiation protocols used in the publications of this thesis we have shown the absence of full-length NFL in human motor neurons upon a nonsense mutation. With the isogenic NFL-KO cell lines we could show that also N-terminal deletions lead to undetectable full-length protein.

Three NFL antibodies were used to assess NFL proteins in the iPSC-MNs: N-terminal (binding site lost in isogenic NFL-KO cell lines), C-terminal (binding site lost in patient cell line) and NFL antibody "pan" with the binding site not reported (raised against pig NF) (Figure 12). In WB no 68kDa signal for full-length NFL was detected with any of the above antibodies in KO or patient MN. In immunocytochemical analysis the patient cell line showed no distinguishable signal from background with N- or C-terminal antibodies, but the isogenic KO cell lines showed stronger signal with the C-terminal antibody compared to the patient. The signal was detectable but in comparison to control it was approximately only 1%. The weak signal did not co-localize with other filament antibodies, indicating either unspecific signal or low amounts of a possible truncated protein. The identified loss of NFL was further complemented by SIMOA measurement of NFL from patient serum and IIF1 KO iPSC-MN cultures, which revealed undetectable NFL in patient serum and trace amounts of NFL in KO media (CTR:5ng and IIF1:0.005ng). Since some signal was detected in KO cell lines, we cannot completely exclude the presence of truncated NFL in NFL-KO cell lines. However, no bands were detected in NFL-absent cell lines from a non-RIPA soluble fraction analyzed by WB whereas control NFL was detectable. Furthermore, long exposure revealed high molecular weight bands with all antibodies in control cell lines that are possibly oligomers of NF proteins containing NFL. No high molecular weight bands were seen in the NFL-KO cell lines, indicating that even though there theoretically might be some truncated protein, according to ICC in the NFL-KO, it is not assembled into detectable oligomers. To conclude, the cell lines studied here are absent of NFL and if there is a truncated protein in the NFL-KOs, it is unable to incorporate into oligomers.

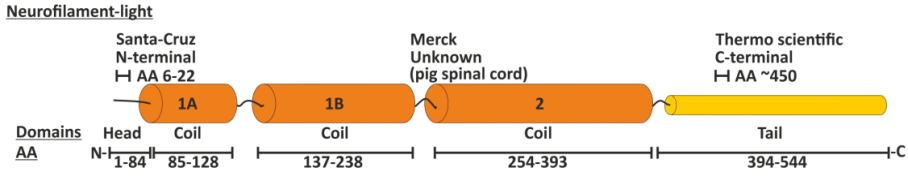

Figure 12. Antibodies against NFL used in III and IV. Amino acids and domains targeted by different NFL antibodies.

Interestingly, N-terminal antibody signal was detected in a sural nerve biopsy of a patient with the p.E210X variant (Yum et al., 2009). However, in transfected cells the p.E210X did not form filaments, indicating a similar situation of NFs devoid of NFL as in our case. Also, similarly as in our model NFs were detected in unmyelinated axons of the p.E210X patient (Yum et al., 2009). In contrast, no protein was detected in another iPSC derived neuronal

NEFL KO model investigating the allele-specific removal of a dominant allele by genetic editing (Feliciano et al., 2021). However, in that study only one antibody was used to assay NFL in WB and no ICC images of the full KO clone were provided.

In literature the term "KO" is in my opinion used too lightly. It is not sufficient to term a cell line as a KO with one antibody in Western blot. Multiple antibodies detecting different epitopes should be used in a cell line with high control expression of the gene, for example NFL KO cannot be evaluated in a non-neural cell, to detect even a small amount that could affect the results. Nonsense mutations can have basal readthrough (Atanasova et al., 2017) or they can be circumvented by exon skipping, producing truncated proteins not detected with all antibodies (Tuladhar et al., 2019).

5.3.3 Differentiation of motor neurons or growth of neurites is not affected by NFL loss (III, IV)

Loss of NFL protein did not reduce the MN differentiation potential of human iPSC, which we assayed with two differentiation protocols. Immunocytochemical analysis of iPSC-MNs lacking NFL showed equal amounts of ISL1/2 positive nuclei in patient, control, and KO cell lines. Significant reduction in differentiation potential is not expected since *NEFL* nonsense mutations do not cause a congenital disease.

The formation of an elaborate neurite network, initial neurite growth (M. S. et. al., unpublished results, Incucyte analysis) or the growth of axons is not impaired in NFL loss. In the most recent and precise analysis with axon isolation devices (Xona) we could show efficient axonal growth and regrowth after axotomy in iPSC-MNs lacking NFL. The axon isolation devices allow the separation of neuronal somas from axons with microgrooves.

Since the structural role of NFL in the NFs and axons is established we were baffled why the axons lacking NFL grow similarly as controls (Yuan et al., 2017). NFL is highly expressed early in neuronal development suggesting it to be integral in neuron differentiation and axon growth (Kirkcaldie & Dwyer, 2017; Pachter & Liem, 1984). One explanation is that, as in mice (Zhu et al., 1997), young human neurons utilize other intermediate filament proteins, such as INA and PRPH, in NFs and can potentially compensate for the loss of NFL. No protein level compensation was detected, but possibly an excess of filament proteins is produced in the maturing iPSC-MNs, which is sufficient for NFs even with no compensation. Moreover, neurons *in vitro* are not subject to mechanical stress that could *in vivo* reduce the regenerative potential of motor neuron axons lacking NFL.

Human patients with nonsense variants in *NEFL* have normal, or mildly affected, growth and development, clearly showing NFL is not needed in initial neuron and axon growth (Abe et al., 2009; Fu & Yuan, 2018; Sainio et al., 2018; Yum et al., 2009). Nevertheless, the patients show a loss of large myelinated axons, indicating NFL and NFs to be necessary in the larger axons in humans (Fu & Yuan, 2018; Yum et al., 2009). Other iPSC-based

studies on NFL that have focused on dominant mutations, which cause NF protein aggregation, have reported no issues in differentiation or axon growth (Saporta et al., 2015; Van Lent et al., 2021). Since nonsense and missense mutations in *NEFL* cause length-depended progressive neuropathy, we can speculate NFL to have a more integral part in NF formation and maintenance in long peripheral motor neuron axons of adults. An *in vitro* co-culture system of motor neurons and Schwann cells allowing long maturation could be used to discover the different requirements for NFL and NF between large myelinated and smaller unmyelinated axons.

5.3.3.1 Variation in differentiation protocols (III, IV)

I experimented with three neuronal differentiation protocols in my thesis. 1) Mixed neuronal differentiation without small molecule inhibitors (Hämäläinen et al., 2013), 2) adherent iPSC-MN differentiation protocol with small molecule inhibitors (Du et al., 2015) and 3) suspension-based iPSC-MN differentiation with small molecule inhibitors (Guo et al., 2017). The first protocol produced a highly heterogenous culture containing multiple types of neurons, glia, and proliferating cells of nondescript origin. Therefore, the first protocol was not utilized further in the study of motor neuron specific axonal degeneration (M. S. et. al., unpublished results).

The second protocol used in the first NFL study produced mainly iPSC neurons and some iPSC-MNs. The adherent protocol with three passaging steps was very laboursome and produced neuronal cultures with very high variability. The confluency and proliferation speed of iPSC in the initial stages of the differentiation had a huge impact in the final product. Too low confluency resulted in reduced proliferation and reduced the yield of mature iPSC-MNs and too high of a confluency resulted in acidified media and apoptosis of cells during adherent culture. Moreover, we had technical difficulties in the short suspension phase in the adherent differentiation. High number of cells died in dissociation before suspension culture, the cells did not recover in suspension (again confluency issues), and the viability of the dissociated spheres was often low resulting in a low yield of neurons, which were overtaken by nondifferentiated proliferating cells upon long culture periods. After the first publication the usage of the protocol was discontinued (Sainio et al., 2018) (M. S. et. al., unpublished results).

The third protocol used was a suspension-based protocol already used in disease modelling and modified (Guo et al., 2017; Maury et al., 2015). The protocol was easy to work with and produced a high number of cells in a shorter time. Some optimizations on iPSC dissociation into suspension, starting cell amount and sphere dissociation resulted in a protocol producing high purity neuronal cultures with up to 95% of ISL1 positive iPSC-MNs. Unfortunately, even this protocol had issues, mainly with the persistence of a small population of proliferating cells overcoming the dish in long cultures. We reasoned the usage of nucleotide analog cancer drugs, such as cytosine arabinoside (AraC) and 5-Fluoro-

2'-deoxyuridine (FdUrd), to be unsuitable for our iPSC-MNs since many studies in the lab are focused on mitochondrial mutations possibly strongly affected by the treatment. The drugs might lead to drug specific, not mutation specific changes. In hindsight, the usage of such drugs could have been beneficial. It is maybe better to have a more pure culture with additional drug effect than a variably impure culture masking iPSC-MN specific changes.

Unfortunately, there is yet no gold standard for iPSC-MN differentiation and many publications lack sufficient detail for protocol reproduction.

5.3.4 Residual *NEFL* mRNA in patient and KO motor neurons (III, IV)

5.3.4.1 mRNA is degraded through nonsense mediated decay (III)

Initial patient derived neurons with *NEFL* nonsense mutation showed a residual 10% of *NEFL* mRNA expression. We showed that the nonsense mutant mRNA was degraded by nonsense mediated decay (NMD), because treatment with CHX, which blocks translation and NMD, increased *NEFL* mRNA amount in patient cells.

An interesting aspect was seen in the control cells in which treatment with CHX reduced *NEFL* mRNA expression, indicating regulation of *NEFL* expression upon translation stalling. This is reasonable since the *NEFL* mRNA and NFL protein are produced in very high quantities, and NFs and their stoichiometric composition is important for motor neuron function (Kirkcaldie & Dwyer, 2017). Without efficient regulation stalled translation could lead to overproduction of *NEFL* mRNA.

5.6.4.2 Residual mRNA is evenly distributed in patient and KO motor neurons (IV)

We saw approximately 10% of residual *NEFL* mRNA also in the isogenic NFL-KO iPSC-MNs with RT-qPCR. It is unexpected that the CRISPR KO cell lines express *NEFL* mRNA on the same level as the nonsense patient's cells. The mutations leading to nonsense mediated mRNA decay are different between patient and CRISPR KO, and in addition the KO lines are missing the DNA in the end part of *NEFL* 5'UTR and the first half of exon 1, while leaving the promoter intact. Sequencing of the residual KO transcripts revealed the mRNA to lack most of exon 1 as expected from the genome editing, but still being transcribed.

We investigated the *NEFL* mRNA localization with in-situ hybridization, to investigate could transported *NEFL* mRNA be protected from NMD. In-situ hybridization allows the fluorescent detection of single mRNA molecules in fixed samples. We did not see a difference in *NEFL* mRNA localization between iPSC-MNs with no NFL and control (Figure 13A). mRNA molecules were evenly reduced in the whole neuron to approximately 15% compared to control (Figure 13B). The *NEFL* mRNA molecules were detected in the cell bodies and neurites (Figure 13C). RNA localization to axons could not be quantified since axonal recognition was not reliable in the imaging, and thus the signal outside the soma was determined as neurite signal. We suspected that the mRNA could be protected during anterograde transport to the axons by miRNAs and RNA-binding proteins (RBP) to

be locally translated (Dalla Costa et al., 2021; Hawley et al., 2019), but no difference was seen in *NEFL* mRNA localization. The transport of mRNA to distal axons and local translation has been established before (Alami et al., 2014).

It is still possible that the detected mRNA molecules are protected by miRNAs or RBPs, but since we saw no accumulation of mRNA in neurites, also non-transported mRNA could be "stored" as backup to be later translated. Additionally, from the isogenic KO cell lines with 5'truncated *NEFL* mRNA we can suggest that the 5'domain of the mRNA is not the binding site for translation inhibiting miRNAs or RBPs.

Could there be a pool of *NEFL* mRNA protected from degradation in human motor neurons ready to be translated upon axonal stress or disruption? RBP immunoprecipitation could be used to sequence the bound mRNA molecules to detect the levels of bound *NEFL* mRNA.

No *NEFL* mRNA was detected in a study of transcription factor induced differentiation of *NEFL* KO iPSC. The authors generated *NEFL* KO by introducing an indel in the first exon of the gene with CRISPR-Cas9. *NEFL* mRNA was possibly not detected since *NEFL* expression was on a motor neuron progenitor level since the cells were only differentiated for 7 days before RNA extraction (Feliciano et al., 2021). However, a proper comparison between KO lines would demand the usage of a similar differentiation strategy. In future KO studies if mRNA-less cell line is desired, a larger genomic cut containing the whole 5'UTR and promoter should be utilized.

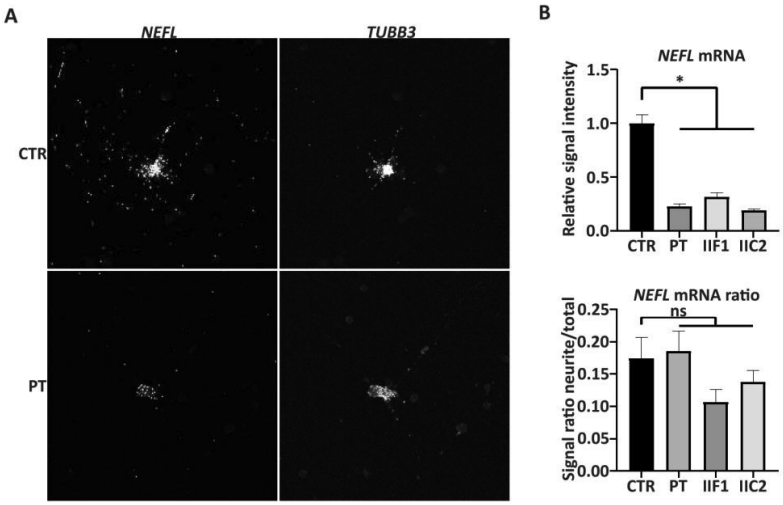

Figure 13. *NEFL* mRNA localization. A. Reprensentative images of *NEFL* and *TUBB3* mRNA localization in iPSC-MNs of control and patient cell lines. B. Total *NEFL* puncta number is reduced in patient and KO iPSC-MNs. C. *NEFL* signal disribution is not altered by nonsense mutations. Proportions of nuclear and neurite *NEFL* signal are the same in all cell lines. * >0.001, one-way ANOVA with multiple comparisons vs CTR.

5.3.5 Neurofilament structure is altered in absence of NFL (III, IV)

The absence of full-length NFL prompted us to investigate whether other filament subunits compensated for its loss and if we could see NFs in the iPSC-MNs lacking NFL. Electron microscopic analysis of iPSC-MNs lacking NFL showed the presence of structures reminiscent of NFs. In addition, immunocytochemistry revealed neurites containing filamentous structures that were positive for all other investigated NF proteins, INA, PRPH, NFM and NFH. Therefore, we are confident that the fetal like iPSC-MNs can form NFs without full-length NFL. NFs are present, although reduced in number, in unmyelinated axons of *NEFL* KO mice (Zhu et al., 1997), as well as in unmyelinated axons of human patients with nonsense *NEFL* variants (Yum et al., 2009).

NF protein levels were not increased in compensation to NFL loss, but NFH levels were reduced indicating changes in the stoichiometry of NFs without NFL. The loss of NFL from NFs possibly causes changes in NF assembly, causing the degradation of unassembled NFH. Similarly, no compensation of other filament proteins and a reduction in NFH levels was seen in the *Nefl* KO mouse (Yuan et al., 2003; Zhu et al., 1997).

We have studied small neurons in culture, with no stress from body movements. Since NFs without NFL and with reduced NFH are able to structurally support the young iPSC-MNs, it would be interesting to grow longer axons in culture to see if the long NFL-less cells are more prone to degeneration. In addition, it could be informative to analyze the mechanical resilience of axons lacking NFL (Grevesse et al., 2015).

5.3.6 Reduced axonal diameter in NFL loss (III, IV)

In contrast to our analysis of mixed iPSC neurons and iPSC-MN neurites in our first study on NFL-loss, we could show the reduction of axonal caliber in our refined human iPSC-MN model. In the initial experiment I investigated the size of neurites, dendrites and axons in a culture containing other neurons in addition to iPSC-MNs. In the more refined differentiation system, I plated iPSC-MNs on drops of laminin to concentrate cell somas on small areas on a plate (Figure 14A). The concentrated iPSC-MNs project axons away from the cluster (Figure 14B). I selected the axonal areas for EM and analyzed the size of the axons.

Gross axonal structure was similar in control and NFL-absent iPSC-MNs. We saw a small, but significant reduction in axon area in all iPSC-MN axons lacking NFL. This is in line with previously reported rodent models lacking NFL that show a reduction in axonal caliber (Jacomy et al., 1999; O. Ohara et al., 1993; Rao et al., 2012; Yuan et al., 2017; Zhu et al., 1997). In the mice, more severe phenotypes were seen in myelinated axons than in unmyelinated. The myelinated axons tended to have no NFs in *Nefl* KO murine models, and the myelinated axon size was reduced more dramatically. In our iPSC-MN model the unmyelinated axons with no NFL tended to have less NFs. The patient and IIC2 KO cell lines had a significantly reduced number of NF-positive axons. The number of NFs in axons could not be reliable counted because of suboptimal contrast in the EM images. The

differences between the first study and the latter exemplify the need for accurate and specific analytical methods to detect small, but significant differences in cell models. The refined model recapitulates the effects of NFL loss seen *in vivo.*

Our *in vitro* findings are comparable with clinical data on nonsense *NEFL* variants (Abe et al., 2009; Fu & Yuan, 2018; Sainio et al., 2018; Yum et al., 2009). All nonsense patients display reduced NCV and the nerve biopsies show reduced axonal caliber and loss of large myelinated axons (Fu & Yuan, 2018; Yum et al., 2009). No demyelination has been seen in patient biopsies, indicating the NCV reduction to stem from decreased diameter of the axons. Therefore, mature human peripheral axons need NFL and functional NFs to maintain a proper caliber and conduction velocity.

It would be informative to further image and analyze the NFL-KO axons to determine the number of NFs and microtubules. Microtubules are likely also affected by the reduction in NFs since NF proteins regulate microtubule polymerization (Yadav et al., 2016).

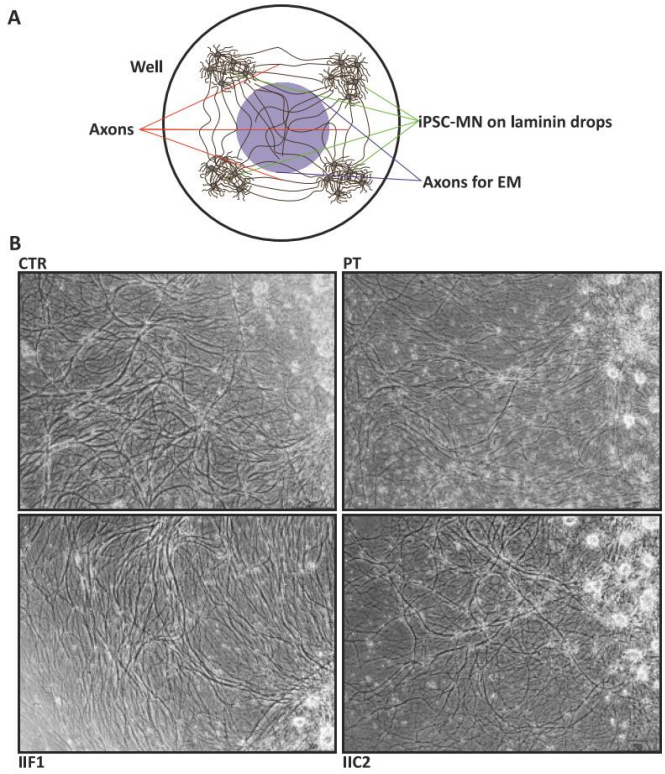

Figure 14. Schematic of axon TEM setup. A. iPSC-MN progenitors are plated on a dry coverslip as a drop to concentrate cell nuclei. Axons extend from the cell nuclei concentrated areas. B. Representative images of iPSC-MN axons next to the nuclei concentrated regions.

5.3.7 Increased movement of organelles in axons devoid of NFL (IV)

Loss of NFL in human iPSC-MN led to increased movement of mitochondria and to lesser extent lysosomes, while the number of organelles within the axons remained the same. I used organelle specific dyes, Mitotracker and Lysotracker, to image the movement of these organelles in live cells. I combined the use of tracker with the microfluidic device to image specifically movement in axons. Analysis of the videos containing hundreds to thousands of organelles revealed the overall movement, anterograde and retrograde combined, to be increased but the speed of movement not to be altered. This indicates that the organelles move longer uninterrupted distances with the same speed.

Since the bulk analysis of organelle movement with fluorescent dyes is subject to bias from analysis artefacts and fluorescent signal intensity, I analyzed selected axons in the videos also as kymographs. I selected clear axons with at least one visibly moving organelle to accurately assay the movement characteristics. The results were comparable with the bulk data for mitochondria, but lysosome movement increase was not seen in kymographs.

Previous studies have shown missense *NEFL* mutations to lead to reduction in mitochondria movement (Saporta et al., 2015; Van Lent et al., 2021). The aggregating NFs and NFL possibly disrupt the trafficking machinery, resulting in movement reduction. In the absence of NFL and aggregates no such disruption happens. In a study of *Nefl* KO mouse dorsal root ganglia neurons the movement of mitochondria and lysosome was also increased (Perrot & Julien, 2009). The increased mitochondrial movement could reduce the amount of stationary active mitochondria in critical axonal locations. The loss of integral mitochondrial functions, such as calcium buffering and oxidative phosphorylation, in high energy demand axonal regions could lead to local energy and metabolic deficiencies (Smith & Gallo, 2018). Collectively the results show NFs to have a regulatory role in organelle trafficking in axons. In cases of both, the dominant aggregating NFL and nonsense-mediated loss of NFL, correct mitochondrial localization can be affected resulting in similar axonal deficiencies.

Length-dependent neuropathies are hypothesized to be caused by organelle trafficking deficiencies (Markworth et al., 2021). It is an easily understandable theory that the long axons would be susceptible to disturbances in organelle movement because the axon tips are far from the nucleus, have high energy demands and are part of a post-mitotic cell. However, further evidence is needed to establish the definitive role of axonal transport in axonal degeneration.

Size analysis of mitochondria by fluorescent imaging provides only a rough estimate. From the tracking experiment I analyzed the average length of mitochondria in axons. The average length of mitochondria was reduced in the IIF1 KO but not in the patient or other KO (M. S. et. al., unpublished results). A detailed EM analysis of mitochondrial structure should be executed to evaluate differences caused by NFL loss. Furthermore, in some studies mutations in *NEFL* have been shown to cause reduction in respiratory chain function (Van Lent et al., 2021), hence it would be interesting to study mitochondrial respiration in NFL

KO iPSC-MNs. Future studies should also address the localization of mitochondria in loss of NFL. Since NFL has been shown to have a synaptic function (Yuan et al., 2018), could it also work as an anchor via Myo Va for mitochondria in the synapses and thus could NFL loss cause a reduction in synaptic mitochondria?

5.3.8 Reduced amplitude of miniature excitatory events in NFL loss (IV)

The electrophysiological function of the differentiated iPSC-MNs was assessed by whole-cell patch-clamp. Week 7 iPSC-MNs of all genotypes showed similar electrophysiological signatures of a mature iPSC-MN. The neurons fired spontaneous and evoked action potentials with similar properties in current-clamp, as well as retained a similar negative membrane potential.

By looking at AMPA receptor mediated currents we could show the amplitude of miniature excitatory post-synaptic currents (mEPSC) to be reduced in all iPSC-MNs lacking NFL. The frequency of events was not affected and is subject to the amount of innervation to the clamped neuron. Recent study (Yuan et al., 2018) showed reduction in post-synaptic AMPA receptors in *Nefl* KO mice. They showed NFL to interact with AMPA receptors possibly protecting them from degradation. The loss of NFL in synapses therefore lead to the reduction of receptors. The reduction of receptors in the post-synaptic cell, seen in mice, could explain the decreased response to released neurotransmitter in our model.

Further experimental evidence is needed to confirm the reduction of receptors on the post-synaptic cell. The reduced amplitude could also result from the pre-synaptic cell in our iPSC-MN cultures. Since NFL has been shown to interact with the ER (Rao et al., 2011) and other organelles (Kim et al., 2011), its loss could lead to disturbances in synaptic vesicle formation and trafficking. The released synaptic vesicles could contain less neurotransmitter, which would be seen as a reduction in amplitude.

Theoretically in a human disorder context reduced amplitude in a post-synaptic lower motor neuron could result in long-term depression, reduced activity, and axonal degeneration in an activating circuit (upper motor neuron to lower motor neuron) and the opposite in an inhibitory circuit (upper motor neuron to inhibitory interneuron to lower motor neuron). In the spinal motor circuit, the lower motor and interneurons post-synaptic terminals could have reduced number of receptors. Thus, an upper-motor neuron signal would less likely induce an action potential in the respective neuron, causing dysregulated activity and muscle activation.

5.3.9 No effect on NFL in patient motor neurons treated with NMDi (IV)

We discovered that a novel *NEFL* nonsense mutation leads to early-onset neuropathy, and that NEFL mRNA is subject to ribosome associated NMD (III). By blocking NMD or increasing the likelihood of translation readthrough by the ribosome we could theoretically induce the production of full-length NFL.

I investigated three such drugs shown in previous studies to induce full-length protein in nonsense cell models (Atanasova et al., 2017; Banning et al., 2018): amlexanox (AMX), ataluren (PTC-124) and gentamicin. No previous nonsense suppression studies have been performed in patient iPSC-MNs. I showed the nonsense suppression drugs to be toxic to developing iPSC-MNs. This is not surprising since the drugs have off-target toxicity (laurent et al., 1982), and insufficient NMD leads to neuron degeneration (Alrahbeni et al., 2015). AMX and gentamicin were toxic at low concentrations and caused increased neuronal disruption earlier. PTC-124 was more tolerated. None of the drugs increased full-length NFL to detectable levels.

Nonsense stop-codon, nucleotide after the stop, basal residual mRNA and basal readthrough affect the likelihood of inducing translational readthrough (Atanasova et al., 2017; Nagel-Wolfrum et al., 2016). In the *NEFL* patient case thymine after the stop-codon and no detectable basal protein predicts no induction of readthrough. On the other hand, the patient cells retained a relative high amount of *NEFL* mRNA, and the UGA stop-codon is predicted to be read through more than other stop codons (Martins-Dias & Romão, 2021). The residual 10% is a large amount of mRNA since the normal expression is very high. Nonetheless, no full-length or truncated protein was detected with the nonsense suppression drugs using the experimented concentrations and treatment times.

In future experiments with neurons a later application of the drugs and experimenting with different next-generation translational readthrough inducing drugs (TRIDs), with reduced off-target toxicity, could induce the readthrough of premature stop codons without toxicity. The clinical translatability of TRIDs has not yet been shown with any nonsense mutation. Studies in cystic fibrosis (Wilschanski et al., 2003), epidermolysis bullosa (Atanasova et al., 2017) and Duchenne muscular dystrophy (Malik et al., 2010) have given promising results in clinical trials, but a positive treatment outcome in patients is yet to be shown.

5.3.10 Axonal neurofilament dysregulation as a common mechanism in CMT

In this thesis I have discovered alterations in iPSC-MNs lacking NFL, reduced mEPSC amplitude, decreased axon size, increased organelle movement and NF dysregulation. NF dysregulation has been also previously suggested to cause a myriad of deficiencies in axons. Changes in NF composition and reduction in the axons could lead to loss of structural support, reduced anchoring scaffolds for organelles, smaller axonal caliber and loss of synaptic scaffolds (Bomont, 2021). Therefore, an underlying common pathological mechanism could be responsible for neuropathy both in NFL loss, studied in this thesis, and in its aggregation via multiple different mechanisms such as loss of NFL protein (Yum et al., 2009), NFL aggregation in the soma (Sasaki et al., 2006) or disrupted NF transport (Brownlees et al., 2002). Furthermore, aggregation in other neuronal disorders could result in similar deficiencies as in the NFL neuropathy (Figure 15) (Balastik et al., 2008; Bomont et al., 2000; Didonna & Opal, 2019; Evgrafov et al., 2004; Zhang et al., 1997). NFL nonsense variants lead to direct loss of NFL in axons and NFL aggregation sequesters NFL

away from the axons. In the case of aggregation, additional pathological effects are at play. Thus, both lead to reduced NFL in axons. Hopefully future studies will investigate the role of NF regulation in these diseases to broaden the patient population that could be targeted by similar treatments.

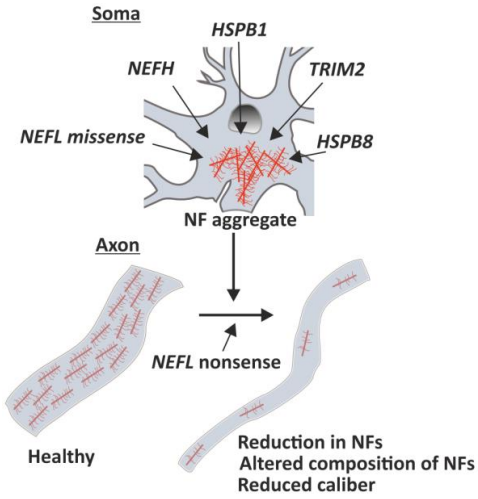

Figure 15. Axonal neurofilament dysregulation as a common mechanism in CMT. Mutations in NF genes and NF folding or degradation regulating genes cause NF aggregation in the soma. Aggregation reduces the amount of correctly folded NFs in axons. *Nefl* KO in mice also leads to reduced NFs in axons.

6 CONCLUSIONS AND FUTURE PERSPECTIVES

The path from a patient with a hereditary monogenic neuromuscular disease into possible treatments is a long and arduous one. My thesis discusses the initial steps in this process, the detection of the causative genetic variant and its modelling *in vitro* with the cell type affected in the disease. Current NGS technologies and variant databases have streamlined the detection of possible pathogenic variants in patients suspected of having a hereditary disease. This has led to the exponential increase of likely pathogenic variants with no functional data or disease modelling. Thus, there is a bottleneck in understanding the pathogenic mechanisms. We have endeavored to investigate the pathogenic mechanisms in two novel variants in known pathogenic genes: *PYROXD1* and *NEFL.*

We showed clinical WES to be an efficient tool in variant detection in a difficult to diagnose adult patient cohort of suspected hereditary neuromuscular and neurological disease. We discovered pathogenic variants in fourteen patients and new likely pathogenic variants in 13 patients, resulting in a finding in 27% of patients with varied neuromuscular disorders. Thus, we expanded the genetic and clinical landscape of neuromuscular disease in Finland. We showed a higher diagnostic yield in patients with an earlier onset and suggest WES to be a cost-effective first-line diagnostic tool in diagnosis of hereditary neuromuscular and neurological disease in adults. Additionally, we determined the molecular diagnosis of a few patients who had been investigated for decades with no conclusive diagnosis, thus ending long diagnostic odysseys.

Clinically the detection of Finnish patients with an uncharacteristic phenotype in *PYROXD1* myopathy is important for the future recognition of more patients with varied phenotypes. It is also important to recognize that the common p.N155S variant with another missense variant can cause an even milder phenotype. Our findings that expanded the phenotypic spectrum of *PYROXD1* myopathy indicate that *PYROXD1* should not only be suspected in congenital myopathy but also in patients with late-onset LGMD. The recognition of *PYROXD1* LGMD is also important for patient care since respiratory insufficiency is seen in all elderly Finnish patients, and should thus be actively monitored. Furthermore, we have shown the p.Y354C variant in *PYROXD1* to reduce the proteins oxidoreductase capacity in yeast challenged with hydrogen peroxide. A similar reduction has been seen with the most common pathogenic missense variant in the gene. However, no studies have been performed in human muscle cells that are the cells affected in the myopathy. Thus, it is not known what is the primarily affected pathway of the reduced oxidoreductase capacity that causes the pathology. The oxidoreductase capacity should be investigated in the affected tissue and antioxidant treatment could be attempted for rescue.

A more comprehensive modelling was performed on the novel p.R367X *NEFL* variant causing early-onset CMT with iPSC-MNs. Our initial analysis suggests the C-terminal nonsense variants in *NEFL* to result in mRNA degradation through nonsense mediated decay and in an absence of NFL protein. Our work on the iPSC derived NFL models

exemplified the need for accurate iPSC-MN differentiation and isogenic controls to model disease mechanisms reliably. Ultimately, we discovered surprisingly unaffected differentiation and growth in addition to possible pathogenic processes in the iPSC-MNs lacking NFL. The loss of an integral NF protein, NFL, does not lead to absence of NFs. Furthermore, iPSC-MN differentiation, axonal growth and basic electrophysiological functions are not hampered in iPSC-MNs devoid of NFL. However, axonal caliber and NFH levels are reduced, mitochondria movement is increased, and the amplitude of miniature excitatory post-synaptic currents is decreased in iPSC-MN lacking NFL. Firstly, reduced axonal size and altered NF composition could reduce the viscoelastic properties and decrease the nerve conduction velocity in patient motor neurons axons. Secondly, increased mitochondrial movement could reduce the amount of stationary active mitochondria in highly metabolically active axonal sites, such as synapses, and result in local metabolic deficiencies. Third, reduced amplitude of miniature post-synaptic currents could lead to decreased neuronal activity upon stimulation. Additionally, we investigated the usability of nonsense skipping drugs in iPSC-MN revealing no readthrough but toxicity. All the alterations detected in iPSC-MN lacking NFL are small but significant, and possibly would not have been detected without isogenic cell lines and accurate analysis of the tissue affected. Hopefully, further research on *NEFL* CMT will replicate our findings, confirm their role in disease pathogenesis and attempt to ameliorate the neuronal phenotypes.

Currently, it is not financially or technically feasible to develop gene or variant specific disease altering treatments for all rare hereditary diseases. However, there is increasing interest in developing treatments for individual patients, and successful treatments have been developed for small patient groups. Gene specific antisense oligonucleotides have been successfully used to increase survival and dramatically improve life quality in a rare (1-2 in 100 000 (Verhaart et al., 2017)) form of spinal muscular atrophy caused by deletions in the *SMA1* gene. Moreover, readthrough inducing drugs have been approved as a treatment in Europe and have been used in small cohorts of Duchenne muscular dystrophy patients with nonsense variants. The long-term treatment has delayed disease progression, but provided only mild benefit for patients (Michael et al., 2021). To achieve benefits for larger patient groups, treating converging common pathological processes shared in different neuronal disorders is more desirable. NF composition alterations seen in NFL loss could be a converging point for many neurological diseases. In CMT caused by dominant missense *NEFL* variants, and variants altering NF protein degradation, NFs have been shown to aggregate in the cell soma and perikarya. The NF aggregation leads to axonal NF reduction and composition changes, reminiscent to *NEFL* nonsense cases. Furthermore, NF aggregates have been seen in other neurodegenerative diseases, such as Alzheimer's disease, Parkinson's disease, amyotrophic lateral sclerosis, and frontotemporal dementia. In all these diseases an axonal NF disruption could lead to similar cellular pathological processes. Highly speculatively, it could be beneficial to increase the stability of existing NFs to ensure a functional network in the axons. Additionally, in NF aggregating cases

transiently decreased NF expression and increased aggregate clearance could increase neuronal survival.

7 ACKNOWLEDGEMENTS

This work was performed in the Stem Cells and Metabolism Program at the Faculty of Medicine, University of Helsinki during the years 2016-2022. This work was financially supported by the Doctoral Program in Biomedicine, the Finnish Brain foundation, Paulo foundation, Biomedicum Helsinki foundation, University of Helsinki, and Academy of Finland.

I would like to thank my supervisors Prof. Henna Tyynismaa and Dr. Emil Ylikallio for the opportunity to work in an atmosphere of high-quality science. You have always found time to mentor and guide me during my years in the lab. I am grateful for the trust you have given me; I have been free to pursue my own ideas and hypotheses. Occasionally my trust in myself has wavered in these past years, but knowing my supervisors believe in my insights has helped me forward. Especially thank you for your friendship and all the relaxed open discussions over a beer.

I would like to thank all the collaborators who have made the research possible, Ras Trovic, Dr. Mari Auranen, Dr. Johanna Palmio, Juho Aaltio and many more.

I thank my thesis reviewers Associate Professor Mika Martikainen and Docent Šárka Lehtonen for taking the time to read and improve my dissertation, and my opponent Pascale Bomont for the public examination.

Thank you, Docent Pirta Hotulainen, and Dr. Maija Castren for being part of my thesis committee and all the insights during the meetings and after.

The second biggest thanks go to all my lab members – Riitta, Taru, Miya, Sundar, Laura, Rosa, Tiina, Jouni, JJ, Junior, Nirajan, Svetlana K., Svetlana M., Matilda, Jay, Mugen, Elina, Nelli, Veijo and Sebastian (hoping I forgot no one!). A special thanks to my closest work friends – Sandra, Jana and Ruben. There has been a great atmosphere in the lab. I have enjoyed all the scientific and completely unscientific discussions shared in our “cool office”.

I would also like to thank my parents Jarmo and Liisa, my older brothers Sami and Kari and my sister Helena, for all the support during the years. I am thankful for the openminded view of the world adopted from my parents, and for the nudge towards science.

And the most thanks go to my wife Camilla for her endless support back home, without her this study would not have been possible. Thank you for coping with me and giving me hope through all these years of doubt in what I do. I also thank my daughter Siiri for while playing with her be it with dolls or dinosaurs, there has been nothing else, giving me respite from the stresses of life and science.

Helsinki 2022,

Markus Sainio

8 REFERENCES

Abe, A., Numakura, C., Saito, K., Koide, H., Oka, N., Honma, A., Kishikawa, Y., & Hayasaka, K. (2009). Neurofilament light chain polypeptide gene mutations in Charcot-Marie-Tooth disease: Nonsense mutation probably causes a recessive phenotype. *Journal of Human Genetics*, *54*(2), 94–97. https://doi.org/10.1038/jhg.2008.13

Adams, D. R., & Eng, C. M. (2018). Next-Generation Sequencing to Diagnose Suspected Genetic Disorders. *New England Journal of Medicine*, *379*(14), 1353–1362. https://doi.org/10.1056/nejmra1711801

Alami, N. H., Smith, R. B., Carrasco, M. A., Williams, L. A., Winborn, C. S., Han, S. S. W., Kiskinis, E., Winborn, B., Freibaum, B. D., Kanagaraj, A., Clare, A. J., Badders, N. M., Bilican, B., Chaum, E., Chandran, S., Shaw, C. E., Eggan, K. C., Maniatis, T., & Taylor, J. P. (2014). Axonal Transport of TDP-43 mRNA Granules Is Impaired by ALS-Causing Mutations. *Neuron*, *81*(3), 536–543. https://doi.org/10.1016/j.neuron.2013.12.018

Alrahbeni, T., Sartor, F., Anderson, J., Miedzybrodzka, Z., McCaig, C., & Müller, B. (2015). Full UPF3B function is critical for neuronal differentiation of neural stem cells. *Molecular Brain*, *8*(1), 33. https://doi.org/10.1186/s13041-015-0122-1

Altshuler, D. L., Durbin, R. M., Abecasis, G. R., Bentley, D. R., Chakravarti, A., Clark, A. G., Collins, F. S., De La Vega, F. M., Donnelly, P., Egholm, M., Flicek, P., Gabriel, S. B., Gibbs, R. A., Knoppers, B. M., Lander, E. S., Lehrach, H., Mardis, E. R., McVean, G. A., Nickerson, D. A., … Peterson, J. L. (2010). A map of human genome variation from population-scale sequencing. *Nature*, *467*(7319), 1061–1073. https://doi.org/10.1038/nature09534

Anderson, C. J., Bredvik, K., Burstein, S. R., Davis, C., Meadows, S. M., Dash, J., Case, L., Milner, T. A., Kawamata, H., Zuberi, A., Piersigilli, A., Lutz, C., & Manfredi, G. (2019). ALS/FTD mutant CHCHD10 mice reveal a tissue-specific toxic gain-of-function and mitochondrial stress response. *Acta Neuropathologica*, *138*(1), 103. https://doi.org/10.1007/s00401-019-01989-y

Anzalone, A. V., Randolph, P. B., Davis, J. R., Sousa, A. A., Koblan, L. W., Levy, J. M., Chen, P. J., Wilson, C., Newby, G. A., Raguram, A., & Liu, D. R. (2019). Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature*, *576*(7785), 149–157. https://doi.org/10.1038/s41586-019-1711-4

Argyrou, A., & Blanchard, J. S. (2004). Flavoprotein Disulfide Reductases: Advances in Chemistry and Function. *Progress in Nucleic Acid Research and Molecular Biology*, *78*, 89–142. https://doi.org/10.1016/S0079-6603(04)78003-4

Aruta, F., Severi, D., Iovino, A., Spina, E., Barghigiani, M., Ruggiero, L., Iodice, R., Santorelli, F. M., Manganelli, F., & Tozza, S. (2021). Proximal weakness involvement in the first Italian case of Charcot-Marie-Tooth 2CC harboring a novel frameshift variant in NEFH. *Journal of the Peripheral Nervous System*, *26*(2), 231–234. https://doi.org/10.1111/jns.12454

Asanović, I., Strandback, E., Kroupova, A., Pasajlic, D., Meinhart, A., Tsung-Pin, P., Djokovic, N., Anrather, D., Schuetz, T., Suskiewicz, M. J., Sillamaa, S., Köcher, T., Beveridge, R., Nikolic, K., Schleiffer, A., Jinek, M., Hartl, M., Clausen, T., Penninger, J., … Martinez, J. (2021). The oxidoreductase PYROXD1 uses NAD(P)+ as an antioxidant to sustain tRNA ligase activity in pre-tRNA splicing and unfolded protein response. *Molecular Cell*, *81*(12), 2520-2532.e16. https://doi.org/10.1016/j.molcel.2021.04.007

Atanasova, V. S., Jiang, Q., Prisco, M., Gruber, C., Piñón Hofbauer, J., Chen, M., Has, C., Bruckner-Tuderman, L., McGrath, J. A., Uitto, J., & South, A. P. (2017). Amlexanox Enhances Premature Termination Codon Read-Through in COL7A1 and Expression of Full Length Type VII Collagen: Potential Therapy for Recessive Dystrophic Epidermolysis Bullosa. *Journal of Investigative Dermatology*, *137*(9), 1842–1849. https://doi.org/10.1016/j.jid.2017.05.011

Auer-Grumbach, M., Toegel, S., Schabhüttl, M., Weinmann, D., Chiari, C., Bennett, D. L. L. H., Beetz, C., Klein, D., Andersen, P. M. M., Böhme, I., Fink-Puches, R., Gonzalez, M., Harms, M. B. B., Motley, W., Reilly, M. M. M., Renner, W., Rudnik-Schöneborn, S., Schlotter-Weigel, B., Themistocleous, A. C. C., … Senderek, J. (2016). Rare Variants in MME, Encoding Metalloprotease Neprilysin, Are Linked to Late-Onset Autosomal-Dominant Axonal Polyneuropathies. *American Journal of Human Genetics*, *99*(3), 607–623. https://doi.org/10.1016/j.ajhg.2016.07.008

Aula, N., Salomäki, P., Timonen, R., Verheijen, F., Mancini, G., Månsson, J. E., Aula, P., & Peltonen, L. (2000). The spectrum of SLC17A5-gene mutations resulting in free sialic acid-storage diseases indicates some genotype-phenotype correlation. *American Journal of Human Genetics*, *67*(4), 832–840. https://doi.org/10.1086/303077

Auranen, M., Ylikallio, E., Shcherbii, M., Paetau, A., Kiuru-Enari, S., Toppila, J. P., & Tyynismaa, H. (2015). CHCHD10 variant p.(Gly66Val) causes axonal Charcot-Marie-Tooth disease. *Neurology.Genetics*, *1*(1), e1. https://doi.org/10.1212/NXG.0000000000000003 [doi]

Balastik, M., Ferraguti, F., Pires-da Silva, A., Tae, H. L., Alvarez-Bolado, G., Kun, P. L., & Gruss, P. (2008). Deficiency in ubiquitin ligase TRIM2 causes accumulation of neurofilament light chain and neurodegeneration. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(33), 12016–12021. https://doi.org/10.1073/pnas.0802261105

Banning, A., Schiff, M., & Tikkanen, R. (2018). Amlexanox provides a potential therapy for nonsense mutations in the lysosomal storage disorder Aspartylglucosaminuria. *Biochimica et Biophysica Acta - Molecular Basis of Disease*, *1864*(3), 668–675. https://doi.org/10.1016/j.bbadis.2017.12.014

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A., & Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, *315*(5819), 1709–1712. https://doi.org/10.1126/science.1138140

Beijer, D., & Baets, J. (2020). The expanding genetic landscape of hereditary motor

neuropathies. *Brain*. https://doi.org/10.1093/brain/awaa311

Bergström, A., Stringer, C., Hajdinjak, M., Scerri, E. M. L., & Skoglund, P. (2021). Origins of modern human ancestry. *Nature 2021 590:7845*, *590*(7845), 229–237. https://doi.org/10.1038/s41586-021-03244-5

Bernard-Marissal, N., Van Hameren, G., Juneja, M., Pellegrino, C., Louhivuori, L., Bartesaghi, L., Rochat, C., El Mansour, O., Médard, J. J., Croisier, M., MacLachlan, C., Poirot, O., Uhlén, P., Timmerman, V., Tricaud, N., Schneider, B. L., & Chrast, R. (2019). Altered interplay between endoplasmic reticulum and mitochondria in Charcot-Marie-Tooth type 2A neuropathy. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(6), 2328–2337. https://doi.org/10.1073/pnas.1810932116

Bian, X., Lin, P., Li, J., Long, F., Duan, R., Yuan, Q., Li, Y., Gao, F., Gao, S., Wei, S., Li, X., Sun, W., Gong, Y., Yan, C., & Liu, Q. (2018). Whole-Genome Linkage Analysis with Whole-Exome Sequencing Identifies a Novel Frameshift Variant in NEFH in a Chinese Family with Charcot-Marie-Tooth 2: A Novel Variant in NEFH for Charcot-Marie-Tooth 2. *Neurodegenerative Diseases*, *18*(2–3), 74–83. https://doi.org/10.1159/000487754

Bis-Brewer, D. M., Fazal, S., & Züchner, S. (2020). Genetic modifiers and non-Mendelian aspects of CMT. In *Brain Research* (Vol. 1726, p. 146459). Elsevier B.V. https://doi.org/10.1016/j.brainres.2019.146459

Bocquet, A., Berges, R., Frank, R., Robert, P., Peterson, A. C., & Eyer, J. (2009). Neurofilaments bind tubulin and modulate its polymerization. *Journal of Neuroscience*, *29*(35), 11043–11054. https://doi.org/10.1523/JNEUROSCI.1924-09.2009

Boerkoel, C. F., Takashima, H., Garcia, C. A., Olney, R. K., Johnson, J., Berry, K., Russo, P., Kennedy, S., Teebi, A. S., Scavina, M., Williams, L. L., Mancias, P., Butler, I. J., Krajewski, K., Shy, M., & Lupski, J. R. (2002). Charcot-Marie-Tooth disease and related neuropathies: Mutation distribution and genotype-phenotype correlation. *Annals of Neurology*, *51*(2), 190–201. https://doi.org/10.1002/ana.10089

Bolduc, V., Marlow, G., Boycott, K. M., Saleki, K., Inoue, H., Kroon, J., Itakura, M., Robitaille, Y., Parent, L., Baas, F., Mizuta, K., Kamata, N., Richard, I., Linssen, W. H. J. P., Mahjneh, I., de Visser, M., Bashir, R., & Brais, B. (2010). Recessive Mutations in the Putative Calcium-Activated Chloride Channel Anoctamin 5 Cause Proximal LGMD2L and Distal MMD3 Muscular Dystrophies. *American Journal of Human Genetics*, *86*(2), 213–221. https://doi.org/10.1016/j.ajhg.2009.12.013

Bomont, P. (2021). The dazzling rise of neurofilaments: Physiological functions and roles as biomarkers. In *Current Opinion in Cell Biology* (Vol. 68, pp. 181–191). Elsevier Ltd. https://doi.org/10.1016/j.ceb.2020.10.011

Bomont, P., Cavalier, L., Blondeau, F., Hamida, C. Ben, Belal, S., Tazir, M., Demir, E., Topaloglu, H., Korinthenberg, R., Tüysüz, B., Landrieu, P., Hentati, F., & Koenig, M. (2000). The gene encoding gigaxonin, a new member of the cytoskeletal BTB/kelch repeat family, is mutated in giant axonal neuropathy. *Nature Genetics*,

26(3), 370–374. https://doi.org/10.1038/81701

Bonanomi, D., & Pfaff, S. L. (2010). Motor axon pathfinding. *Cold Spring Harbor Perspectives in Biology*, *2*(3). https://doi.org/10.1101/cshperspect.a001735

Boycott, K. M., Vanstone, M. R., Bulman, D. E., & MacKenzie, A. E. (2013). Rare-disease genetics in the era of next-generation sequencing: Discovery to translation. *Nature Reviews Genetics*, *14*(10), 681–691. https://doi.org/10.1038/nrg3555

Bravo-Ambrosio, A., & Kaprielian, Z. (2011). Crossing the border: Molecular control of motor axon exit. *International Journal of Molecular Sciences*, *12*(12), 8539–8561. https://doi.org/10.3390/ijms12128539

Brownlees, J., Ackerley, S., Grierson, A. J., Jacobsen, N. J. O., Shea, K., Anderton, B. H., Leigh, P. N., Shaw, C. E., & Miller, C. C. J. (2002). Charcot-Marie-Tooth disease neurofilament mutations disrupt neurofilament assembly and axonal transport. *Human Molecular Genetics*, *11*(23), 2837–2844. https://doi.org/10.1093/hmg/11.23.2837

Chambers, S. M., Fasano, C. A., Papapetrou, E. P., Tomishima, M., Sadelain, M., & Studer, L. (2009). Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nature Biotechnology*, *27*(3), 275–280. https://doi.org/10.1038/nbt.1529

Chen, Y., & Lai, D. (2015). Pluripotent states of human embryonic stem cells. *Cellular Reprogramming*, *17*(1), 1–6. https://doi.org/10.1089/cell.2014.0061

Chernyatina, A. A., Guzenko, D., & Strelkov, S. V. (2015). Intermediate filament structure: The bottom-up approach. In *Current Opinion in Cell Biology* (Vol. 32, pp. 65–72). Elsevier Ltd. https://doi.org/10.1016/j.ceb.2014.12.007

Chernyatina, A. A., Nicolet, S., Aebi, U., Herrmann, H., & Strelkov, S. V. (2012). Atomic structure of the vimentin central α-helical domain and its implications for intermediate filament assembly. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(34), 13620–13625. https://doi.org/10.1073/pnas.1206836109

Cho, H. H., Cargnin, F., Kim, Y., Lee, B., Kwon, R. J., Nam, H., Shen, R., Barnes, A. P., Lee, J. W., Lee, S., & Lee, S. K. (2014). Isl1 Directly Controls a Cholinergic Neuronal Identity in the Developing Forebrain and Spinal Cord by Forming Cell Type-Specific Complexes. *PLoS Genetics*, *10*(4). https://doi.org/10.1371/journal.pgen.1004280

Chong, J. X., Buckingham, K. J., Jhangiani, S. N., Boehm, C., Sobreira, N., Smith, J. D., Harrell, T. M., McMillin, M. J., Wiszniewski, W., Gambin, T., Coban Akdemir, Z. H., Doheny, K., Scott, A. F., Avramopoulos, D., Chakravarti, A., Hoover-Fong, J., Mathews, D., Witmer, P. D., Ling, H., … Bamshad, M. J. (2015). The Genetic Basis of Mendelian Phenotypes: Discoveries, Challenges, and Opportunities. *American Journal of Human Genetics*, *97*(2), 199–215. https://doi.org/10.1016/j.ajhg.2015.06.009

Córdoba, M., Rodriguez-Quiroga, S. A., Vega, P. A., Salinas, V., Perez-Maturo, J.,

Amartino, H., Vásquez-Dusefante, C., Medina, N., González-Morón, D., & Kauffman, M. A. (2018). Whole exome sequencing in neurogenetic odysseys: An effective, cost- and time-saving diagnostic approach. *PLoS ONE*, *13*(2). https://doi.org/10.1371/journal.pone.0191228

Cortese, A., Zhu, Y., Rebelo, A. P., Negri, S., Courel, S., Abreu, L., Bacon, C. J., Bai, Y., Bis-Brewer, D. M., Bugiardini, E., Buglo, E., Danzi, M. C., Feely, S. M. E., Athanasiou-Fragkouli, A., Haridy, N. A., Rodriguez, A., Bacha, A., Kosikowski, A., Wood, B., … Zuchner, S. (2020). Biallelic mutations in SORD cause a common and potentially treatable hereditary neuropathy with implications for diabetes. *Nature Genetics*, *52*(5), 473–481. https://doi.org/10.1038/s41588-020-0615-4

Daimagüler, H. S., Akpulat, U., Özdemir, Ö., Yis, U., Güngör, S., Talim, B., Diniz, G., Baydan, F., Thiele, H., Altmüller, J., Nürnberg, P., & Cirak, S. (2021). Clinical and genetic characterization of PYROXD1-related myopathy patients from Turkey. *American Journal of Medical Genetics, Part A*, *185*(6), 1678–1690. https://doi.org/10.1002/ajmg.a.62148

Dalla Costa, I., Buchanan, C. N., Zdradzinski, M. D., Sahoo, P. K., Smith, T. P., Thames, E., Kar, A. N., & Twiss, J. L. (2021). The functional organization of axonal mRNA transport and translation. In *Nature Reviews Neuroscience* (Vol. 22, Issue 2, pp. 77–91). Nature Research. https://doi.org/10.1038/s41583-020-00407-7

Dasen, J. S., & Jessell, T. M. (2009). Chapter Six Hox Networks and the Origins of Motor Neuron Diversity. *Current Topics in Developmental Biology*, *88*, 169–200. https://doi.org/10.1016/S0070-2153(09)88006-X

Dasen, J. S., Tice, B. C., Brenner-Morton, S., & Jessell, T. M. (2005). A Hox regulatory network establishes motor neuron pool identity and target-muscle connectivity. *Cell*, *123*(3), 477–491. https://doi.org/10.1016/j.cell.2005.09.009

Davis-Dusenbery, B. N., Williams, L. A., Klim, J. R., & Eggan, K. (2014). How to make spinal motor neurons. *Development (Cambridge)*, *141*(3), 491–501. https://doi.org/10.1242/dev.097410

Delisle, M. B., & Carpenter, S. (1984). Neurofibrillary axonal swellings and amyotrophic lateral sclerosis. *Journal of the Neurological Sciences*, *63*(2), 241–250. https://doi.org/10.1016/0022-510X(84)90199-0

Depondt, C., Donatello, S., Rai, M., Wang, F. C., Manto, M., Simonis, N., & Pandolfo, M. (2016). MME mutation in dominant spinocerebellar ataxia with neuropathy (SCA43). *Neurology: Genetics*, *2*(5), 94. https://doi.org/10.1212/NXG.0000000000000094

Didonna, A., & Opal, P. (2019). The role of neurofilament aggregation in neurodegeneration: Lessons from rare inherited neurological disorders. In *Molecular Neurodegeneration* (Vol. 14, Issue 1, pp. 1–10). BioMed Central Ltd. https://doi.org/10.1186/s13024-019-0318-4

Djebali, S., Davis, C. A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., Tanzer, A., Lagarde, J., Lin, W., Schlesinger, F., Xue, C., Marinov, G. K., Khatun, J., Williams, B. A., Zaleski, C., Rozowsky, J., Röder, M., Kokocinski, F., Abdelhamid, R. F., …

Gingeras, T. R. (2012). Landscape of transcription in human cells. *Nature*, *489*(7414), 101–108. https://doi.org/10.1038/nature11233

Dowling, J. J., Weihl, C. C., & Spencer, M. J. (2021). Molecular and cellular basis of genetically inherited skeletal muscle disorders. *Nature Reviews Molecular Cell Biology*, *22*(11), 713–732. https://doi.org/10.1038/s41580-021-00389-z

Du, Z. W., Chen, H., Liu, H., Lu, J., Qian, K., Huang, C. T. L., Zhong, X., Fan, F., & Zhang, S. C. (2015). Generation and expansion of highly pure motor neuron progenitors from human pluripotent stem cells. *Nature Communications*, *6*. https://doi.org/10.1038/ncomms7626

Dudanova, I., Kao, T. J., Herrmann, J. E., Zheng, B., Kania, A., & Klein, R. (2012). Genetic evidence for a contribution of EphA:EphrinA reverse signaling to motor axon guidance. *Journal of Neuroscience*, *32*(15), 5209–5215. https://doi.org/10.1523/JNEUROSCI.5707-11.2012

Echaniz-Laguna, A., Lornage, X., Lannes, B., Schneider, R., Bierry, G., Dondaine, N., Boland, A., Deleuze, J. F., Böhm, J., Thompson, J., Laporte, J., & Biancalana, V. (2017). HSPB8 haploinsufficiency causes dominant adult-onset axial and distal myopathy. *Acta Neuropathologica*, *134*(1), 163–165. https://doi.org/10.1007/s00401-017-1724-8

Eratne, D., Schneider, A., Lynch, E., Martyn, M., Velakoulis, D., Fahey, M., Kwan, P., Leventer, R., Rafehi, H., Chong, B., Stark, Z., Lunke, S., Phelan, D. G., O'Keefe, M., Siemering, K., West, K., Sexton, A., Jarmolowicz, A., Taylor, J. A., … Berkovic, S. F. (2021). The clinical utility of exome sequencing and extended bioinformatic analyses in adolescents and adults with a broad range of neurological phenotypes: an Australian perspective. *Journal of the Neurological Sciences*, *420*. https://doi.org/10.1016/j.jns.2020.117260

Ericson, J., Thor, S., Edlund, T., Jessell, T. M., & Yamada, T. (1992). Early stages of motor neuron differentiation revealed by expression of homeobox gene Islet-1. *Science*, *256*(5063), 1555–1560. https://doi.org/10.1126/science.1350865

Evgrafov, O. V., Mersiyanova, I., Irobi, J., Van Den Bosch, L. V., Dierick, I., Leung, C. L., Schagina, O., Verpoorten, N., Van Impe, K., Fedotov, V., Dadali, E., Auer-Grumbach, M., Windpassinger, C., Wagner, K., Mitrovic, Z., Hilton-Jones, D., Talbot, K., Martin, J. J., Vasserman, N., … Timmerman, V. (2004). Mutant small heat-shock protein 27 causes axonal Charcot-Marie-Tooth disease and distal hereditary motor neuropathy. *Nature Genetics*, *36*(6), 602–606. https://doi.org/10.1038/ng1354

Ewans, L. J., Schofield, D., Shrestha, R., Zhu, Y., Gayevskiy, V., Ying, K., Walsh, C., Lee, E., Kirk, E. P., Colley, A., Ellaway, C., Turner, A., Mowat, D., Worgan, L., Freckmann, M. L., Lipke, M., Sachdev, R., Miller, D., Field, M., … Roscioli, T. (2018). Whole-exome sequencing reanalysis at 12 months boosts diagnosis and is cost-effective when applied early in Mendelian disorders. *Genetics in Medicine 2018 20:12*, *20*(12), 1564–1574. https://doi.org/10.1038/gim.2018.39

Faber, C. G., Hoeijmakers, J. G. J., Ahn, H. S., Cheng, X., Han, C., Choi, J. S., Estacion,

M., Lauria, G., Vanhoutte, E. K., Gerrits, M. M., Dib-Hajj, S., Drenth, J. P. H., Waxman, S. G., & Merkies, I. S. J. (2012). Gain of function Nav1.7 mutations in idiopathic small fiber neuropathy. *Annals of Neurology*, *71*(1), 26–39. https://doi.org/10.1002/ANA.22485

Fabrizi, G. M., Cavallaro, T., Angiari, C., Cabrini, I., Taioli, F., Malerba, G., Bertolasi, L., & Rizzuto, N. (2007). Charcot-Marie-Tooth disease type 2E, a disorder of the cytoskeleton. *Brain*, *130*(2), 394–403. https://doi.org/10.1093/brain/awl284

Feliciano, C. M., Wu, K., Watry, H. L., Marley, C. B. E., Ramadoss, G. N., Ghanim, H. Y., Liu, A. Z., Zholudeva, L. V., McDevitt, T. C., Saporta, M. A., Conklin, B. R., & Judge, L. M. (2021). Allele-Specific Gene Editing Rescues Pathology in a Human Model of Charcot-Marie-Tooth Disease Type 2E. *Frontiers in Cell and Developmental Biology*, *9*, 2226. https://doi.org/10.3389/fcell.2021.723023

Fernandez-Marmiesse, A., Gouveia, S., & Couce, M. L. (2017). NGS Technologies as a Turning Point in Rare Disease Research , Diagnosis and Treatment. *Current Medicinal Chemistry*, *25*(3), 404–432. https://doi.org/10.2174/0929867324666170718101946

Fernandopulle, M. S., Prestil, R., Grunseich, C., Wang, C., Gan, L., & Ward, M. E. (2018). Transcription-factor mediated differentiation of human iPSCs into neurons. *Current Protocols in Cell Biology*, *79*(1), e51. https://doi.org/10.1002/CPCB.51

Filograna, R., Mennuni, M., Alsina, D., & Larsson, N. G. (2021). Mitochondrial DNA copy number in human disease: the more the better? *FEBS Letters*, *595*(8), 976–1002. https://doi.org/10.1002/1873-3468.14021

Fortun, J., Go, J. C., Li, J., Amici, S. A., Dunn, W. A., & Notterpek, L. (2006). Alterations in degradative pathways and protein aggregation in a neuropathy model based on PMP22 overexpression. *Neurobiology of Disease*, *22*(1), 153–164. https://doi.org/10.1016/j.nbd.2005.10.010

Fu, J., & Yuan, Y. (2018). A novel homozygous nonsense mutation in NEFL causes autosomal recessive Charcot–Marie–Tooth disease. *Neuromuscular Disorders*, *28*(1), 44–47. https://doi.org/10.1016/j.nmd.2017.09.018

Fujimori, K., Ishikawa, M., Otomo, A., Atsuta, N., Nakamura, R., Akiyama, T., Hadano, S., Aoki, M., Saya, H., Sobue, G., & Okano, H. (2018). Modeling sporadic ALS in iPSC-derived motor neurons identifies a potential therapeutic agent. *Nature Medicine*, *24*(10), 1579–1589. https://doi.org/10.1038/s41591-018-0140-5

Gafson, A. R., Barthélemy, N. R., Bomont, P., Carare, R. O., Durham, H. D., Julien, J. P., Kuhle, J., Leppert, D., Nixon, R. A., Weller, R. O., Zetterberg, H., & Matthews, P. M. (2020). Neurofilaments: Neurobiological foundations for biomarker applications. In *Brain* (Vol. 143, Issue 7, pp. 1975–1998). Oxford University Press. https://doi.org/10.1093/brain/awaa098

Garcia, M. L., Lobsiger, C. S., Shah, S. B., Deerinck, T. J., Crum, J., Young, D., Ward, C. M., Crawford, T. O., Gotow, T., Uchiyama, Y., Ellisman, M. H., Calcutt, N. A., & Cleveland, D. W. (2003). NF-M is an essential target for the myelin-directed

"outside-in" signaling cascade that mediates radial axonal growth. *Journal of Cell Biology*, *163*(5), 1011–1020. https://doi.org/10.1083/jcb.200308159

Garcia, N. V. D. M., & Jessell, T. M. (2008). Early Motor Neuron Pool Identity and Muscle Nerve Trajectory Defined by Postmitotic Restrictions in Nkx6.1 Activity. *Neuron*, *57*(2), 217–231. https://doi.org/10.1016/j.neuron.2007.11.033

Gasiunas, G., Barrangou, R., Horvath, P., & Siksnys, V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(39). https://doi.org/10.1073/pnas.1208507109

Geisler, N., & Weber, K. (1982). The amino acid sequence of chicken muscle desmin provides a common structural model for intermediate filament proteins. *The EMBO Journal*, *1*(12), 1649–1656. https://doi.org/10.1002/j.1460-2075.1982.tb01368.x

Gentil, B. J., Lai, G.-T., Menade, M., Larivière, R., Minotti, S., Gehring, K., Chapple, J.-P., Brais, B., & Durham, H. D. (2019). Sacsin, mutated in the ataxia ARSACS, regulates intermediate filament assembly and dynamics. *The FASEB Journal*, *33*(2), 2982–2994. https://doi.org/10.1096/FJ.201801556R

Gentil, B. J., Minotti, S., Beange, M., Baloh, R. H., Julien, J., & Durham, H. D. (2012). Normal role of the low-molecular-weight neurofilament protein in mitochondrial dynamics and disruption in Charcot-Marie-Tooth disease. *The FASEB Journal*, *26*(3), 1194–1203. https://doi.org/10.1096/fj.11-196345

Georganopoulou, D. G., Moisiadis, V. G., Malik, F. A., Mohajer, A., Dashevsky, T. M., Wuu, S. T., & Hu, C. K. (2021). A Journey with LGMD: From Protein Abnormalities to Patient Impact. *Protein Journal*, *40*(4), 466–488. https://doi.org/10.1007/s10930-021-10006-9

Goldman, J. E., Yen, S. H., Chiu, F. C., & Peress, N. S. (1983). Lewy bodies of Parkinson's disease contain neurofilament antigens. *Science*, *221*(4615), 1082–1084. https://doi.org/10.1126/science.6308771

Gorcenco, S., Ilinca, A., Almasoudi, W., Kafantari, E., Lindgren, A. G., & Puschmann, A. (2020). New generation genetic testing entering the clinic. *Parkinsonism and Related Disorders*, *73*, 72–84. https://doi.org/10.1016/j.parkreldis.2020.02.015

Grevesse, T., Dabiri, B. E., Parker, K. K., & Gabriele, S. (2015). Opposite rheological properties of neuronal microcompartments predict axonal vulnerability in brain injury. *Scientific Reports*, *5*. https://doi.org/10.1038/srep09475

Guo, W., Naujock, M., Fumagalli, L., Vandoorne, T., Baatsen, P., Boon, R., Ordovás, L., Patel, A., Welters, M., Vanwelden, T., Geens, N., Tricot, T., Benoy, V., Steyaert, J., Lefebvre-Omar, C., Boesmans, W., Jarpe, M., Sterneckert, J., Wegner, F., … Van Den Bosch, L. (2017). HDAC6 inhibition reverses axonal transport defects in motor neurons derived from FUS-ALS patients. *Nature Communications*, *8*(1). https://doi.org/10.1038/s41467-017-00911-y

Hackman, P., Vihola, A., Haravuori, H., Marchand, S., Sarparanta, J., De Seze, J., Labeit, S., Witt, C., Peltonen, L., Richard, I., & Udd, B. (2002). Tibial muscular dystrophy is

a titinopathy caused by mutations in TTN, the gene encoding the giant skeletal-muscle protein titin. *American Journal of Human Genetics*, *71*(3), 492–500. https://doi.org/10.1086/342380

Haidar, M., Asselbergh, B., Adriaenssens, E., De Winter, V., Timmermans, J. P., Auer-Grumbach, M., Juneja, M., & Timmerman, V. (2019). Neuropathy-causing mutations in HSPB1 impair autophagy by disturbing the formation of SQSTM1/p62 bodies. *Autophagy*, *15*(6), 1051–1068. https://doi.org/10.1080/15548627.2019.1569930

Hakonen, A. H., Davidzon, G., Salemi, R., Bindoff, L. A., Van Goethem, G., DiMauro, S., Thorburn, D. R., & Suomalainen, A. (2007). Abundance of the POLG disease mutations in Europe, Australia, New Zealand, and the United States explained by single ancient European founders. *European Journal of Human Genetics*, *15*(7), 779–783. https://doi.org/10.1038/sj.ejhg.5201831

Hakonen, A. H., Heiskanen, S., Juvonen, V., Lappalainen, I., Luoma, P. T., Rantamäki, M., Van Goethem, G., Löfgren, A., Hackman, P., Paetau, A., Kaakkola, S., Majamaa, K., Varilo, T., Udd, B., Kääriäinen, H., Bindoff, L. A., & Suomalainen, A. (2005). Mitochondrial DNA polymerase W748S mutation: A common cause of autosomal recessive ataxia with ancient European origin. *American Journal of Human Genetics*, *77*(3), 430–441. https://doi.org/10.1086/444548

Hämäläinen, R. H., Manninen, T., Koivumäki, H., Kislin, M., Otonkoski, T., & Suomalainen, A. (2013). Tissue-and cell-type-specific manifestations of heteroplasmic mtDNA 3243A>G mutation in human induced pluripotent stem cell-derived disease model. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(38). https://doi.org/10.1073/pnas.1311660110

Hamburger, V. (1975). Cell death in the development of the lateral motor column of the chick embryo. *Journal of Comparative Neurology*, *160*(4), 535–546. https://doi.org/10.1002/cne.901600408

Hanemann, C. O., Gabreëls-Festen, A. A. W. M., Stoll, G., & Müller, H. W. (1997). Schwann cell differentiation in Charcot-Marie-Tooth disease type 1A (CMT1A): Normal number of myelinating Schwann cells in young CMT1A patients and neural cell adhesion molecule expression in onion bulbs. *Acta Neuropathologica*, *94*(4), 310–315. https://doi.org/10.1007/s004010050712

Harjuhaahto, S., Rasila, T. S., Molchanova, S. M., Woldegebriel, R., Kvist, J., Konovalova, S., Sainio, M. T., Pennonen, J., Torregrosa-Muñumer, R., Ibrahim, H., Otonkoski, T., Taira, T., Ylikallio, E., & Tyynismaa, H. (2020). ALS and Parkinson's disease genes CHCHD10 and CHCHD2 modify synaptic transcriptomes in human iPSC-derived motor neurons. *Neurobiology of Disease*, *141*. https://doi.org/10.1016/j.nbd.2020.104940

Hartley, T., Wagner, J. D., Warman-Chardon, J., Tetreault, M., Brady, L., Baker, S., Tarnopolsky, M., Bourque, P. R., Parboosingh, J. S., Smith, C., McInnes, B., Innes, A. M., Bernier, F., Curry, C. J., Yoon, G., Horvath, G. A., Bareke, E., Gillespie, M., Consortium, F. C., … Boycott, K. M. (2018). Whole-exome sequencing is a valuable diagnostic tool for inherited peripheral neuropathies: Outcomes from a cohort of 50

families. *Clinical Genetics*, *93*(2), 301–309. https://doi.org/10.1111/cge.13101 [doi]

Hawley, Z. C. E., Campos-Melo, D., & Strong, M. J. (2019). MiR-105 and miR-9 regulate the mRNA stability of neuronal intermediate filaments. Implications for the pathogenesis of amyotrophic lateral sclerosis (ALS). *Brain Research*, *1706*, 93–100. https://doi.org/10.1016/j.brainres.2018.10.032

Heather, J. M., & Chain, B. (2016). The sequence of sequencers: The history of sequencing DNA. In *Genomics* (Vol. 107, Issue 1, pp. 1–8). Academic Press Inc. https://doi.org/10.1016/j.ygeno.2015.11.003

Hellberg, C., Alinder, E., Jaraj, D., & Puschmann, A. (2019). Nationwide prevalence of primary dystonia, progressive ataxia and hereditary spastic paraplegia. *Parkinsonism and Related Disorders*, *69*, 79–84. https://doi.org/10.1016/j.parkreldis.2019.10.028

Henn, B. M., Botigué, L. R., Bustamante, C. D., Clark, A. G., & Gravel, S. (2015). Estimating the mutation load in human genomes. *Nature Reviews Genetics*, *16*(6), 333–343. https://doi.org/10.1038/nrg3931

Herrmann, H., & Aebi, U. (2016). Intermediate filaments: Structure and assembly. *Cold Spring Harbor Perspectives in Biology*, *8*(11). https://doi.org/10.1101/cshperspect.a018242

Herrmann, H., Häner, M., Brettel, M., Müller, S. A., Goldie, K. N., Fedtke, B., Lustig, A., Franke, W. W., & Aebi, U. (1996). Structure and assembly properties of the intermediate filament protein vimentin: The role of its head, rod and tail domains. *Journal of Molecular Biology*, *264*(5), 933–953. https://doi.org/10.1006/jmbi.1996.0688

Herrmann, H., Strelkov, S. V., Burkhard, P., & Aebi, U. (2009). Intermediate filaments: Primary determinants of cell architecture and plasticity. *Journal of Clinical Investigation*, *119*(7), 1772–1783. https://doi.org/10.1172/JCI38214

Hicks, D., Sarkozy, A., Muelas, N., Koehler, K., Huebner, A., Hudson, G., Chinnery, P. F., Barresi, R., Eagle, M., Polvikoski, T., Bailey, G., Miller, J., Radunovic, A., Hughes, P. J., Roberts, R., Krause, S., Walter, M. C., Laval, S. H., Straub, V., … Bushby, K. (2011). A founder mutation in Anoctamin 5 is a major cause of limb-girdle muscular dystrophy. *Brain*, *134*(1), 171–182. https://doi.org/10.1093/brain/awq294

Higuchi, Y., Hashiguchi, A., Yuan, J., Yoshimura, A., Mitsui, J., Ishiura, H., Tanaka, M., Ishihara, S., Tanabe, H., Nozuma, S., Okamoto, Y., Matsuura, E., Ohkubo, R., Inamizu, S., Shiraishi, W., Yamasaki, R., Ohyagi, Y., Kira, J. I., Oya, Y., … Takashima, H. (2016). Mutations in MME cause an autosomal-recessive Charcot-Marie-Tooth disease type 2. *Annals of Neurology*, *79*(4), 659–672. https://doi.org/10.1002/ana.24612

Hisanaga, S. ichi, & Hirokawa, N. (1988). Structure of the peripheral domains of neurofilaments revealed by low angle rotary shadowing. *Journal of Molecular Biology*, *202*(2), 297–305. https://doi.org/10.1016/0022-2836(88)90459-7

Ikenberg, E., Reilich, P., Abicht, A., Heller, C., Schoser, B., & Walter, M. C. (2019).

Charcot-Marie-Tooth disease type 2CC due to a frameshift mutation of the neurofilament heavy polypeptide gene in an Austrian family. *Neuromuscular Disorders*, *29*(5), 392–397. https://doi.org/10.1016/J.NMD.2019.02.007

Ishii, T., Haga, S., & Tokutake, S. (1979). Presence of neurofilament protein in Alzheimer's neurofibrillary tangles (ANT) - An immunofluorescent study. *Acta Neuropathologica*, *48*(2), 105–112. https://doi.org/10.1007/BF00691151

Israeli, E., Dryanovski, D. I., Schumacker, P. T., Chandel, N. S., Singer, J. D., Julien, J. P., Goldman, R. D., & Opal, P. (2016). Intermediate filament aggregates cause mitochondrial dysmotility and increase energy demands in giant axonal neuropathy. *Human Molecular Genetics*, *25*(11), 2143–2157. https://doi.org/10.1093/hmg/ddw081

Jacomy, H., Zhu, Q., Couillard-Després, S., Beaulieu, J. M., & Julien, J. P. (1999). Disruption of type IV intermediate filament network in mice lacking the neurofilament medium and heavy subunits. *Journal of Neurochemistry*, *73*(3), 972–984. https://doi.org/10.1046/j.1471-4159.1999.0730972.x

Jacquier, A., Delorme, C., Belotti, E., Juntas-Morales, R., Solé, G., Dubourg, O., Giroux, M., Maurage, C. A., Castellani, V., Rebelo, A., Abrams, A., Züchner, S., Stojkovic, T., Schaeffer, L., & Latour, P. (2017). Cryptic amyloidogenic elements in mutant NEFH causing Charcot-Marie-Tooth 2 trigger aggresome formation and neuronal death. *Acta Neuropathologica Communications*, *5*(1), 55. https://doi.org/10.1186/s40478-017-0457-1

Jaffer, F., Murphy, S. M., Scoto, M., Healy, E., Rossor, A. M., Brandner, S., Phadke, R., Selcen, D., Jungbluth, H., Muntoni, F., & Reilly, M. M. (2012). BAG3 mutations: Another cause of giant axonal neuropathy. *Journal of the Peripheral Nervous System*, *17*(2), 210–216. https://doi.org/10.1111/j.1529-8027.2012.00409.x

Jessell, T. M. (2000). Neuronal specification in the spinal cord: Inductive signals and transcriptional codes. In *Nature Reviews Genetics* (Vol. 1, Issue 1, pp. 20–29). Nat Rev Genet. https://doi.org/10.1038/35049541

Jha, M. K., & Morrison, B. M. (2018). Glia-neuron energy metabolism in health and diseases: New insights into the role of nervous system metabolic transporters. In *Experimental Neurology* (Vol. 309, pp. 23–31). Academic Press Inc. https://doi.org/10.1016/j.expneurol.2018.07.009

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, *337*(6096), 816–821. https://doi.org/10.1126/science.1225829

Jokela, M., Penttilä, S., Huovinen, S., Hackman, P., Saukkonen, A. M., Toivanen, J., & Udd, B. (2011). Late-onset lower motor neuronopathy: A new autosomal dominant disorder. *Neurology*, *77*(4), 334–340. https://doi.org/10.1212/WNL.0b013e3182267b71

Jordan, L. M., McVagh, J. R., Noga, B. R., Cabaj, A. M., Majczyński, H., Sławińska, U., Provencher, J., Leblond, H., & Rossignol, S. (2014). Cholinergic mechanisms in

spinal locomotion — Potential target for rehabilitation approaches. *Frontiers in Neural Circuits*, *8*, 132. https://doi.org/10.3389/FNCIR.2014.00132/BIBTEX

Juneja, M., Azmi, A., Baets, J., Roos, A., Jennings, M. J., Saveri, P., Pisciotta, C., Bernard-Marissal, N., Schneider, B. L., Verfaillie, C., Chrast, R., Seeman, P., Hahn, A. F., De Jonghe, P., Maudsley, S., Horvath, R., Pareyson, D., & Timmerman, V. (2018). PFN2 and GAMT as common molecular determinants of axonal Charcot-Marie-Tooth disease. *Journal of Neurology, Neurosurgery and Psychiatry*, *89*(8), 870–878. https://doi.org/10.1136/jnnp-2017-317562

Juneja, M., Burns, J., Saporta, M. A., & Timmerman, V. (2019). Challenges in modelling the Charcot-Marie-Tooth neuropathies for therapy development. *Journal of Neurology, Neurosurgery and Psychiatry*, *90*(1), 58–67. https://doi.org/10.1136/jnnp-2018-318834

Kantor, A., McClements, M. E., & Maclaren, R. E. (2020). Crispr-cas9 dna base-editing and prime-editing. *International Journal of Molecular Sciences*, *21*(17), 1–22. https://doi.org/10.3390/ijms21176240

Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alföldi, J., Wang, Q., Collins, R. L., Laricchia, K. M., Ganna, A., Birnbaum, D. P., Gauthier, L. D., Brand, H., Solomonson, M., Watts, N. A., Rhodes, D., Singer-Berk, M., England, E. M., Seaby, E. G., Kosmicki, J. A., … MacArthur, D. G. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*, *581*(7809), 434–443. https://doi.org/10.1038/s41586-020-2308-7

Kenvin, S., Torregrosa-Muñumer, R., Reidelbach, M., Pennonen, J., Turkia, J. J., Rannila, E., Kvist, J., Sainio, M. T., Huber, N., Herukka, S.-K., Haapasalo, A., Auranen, M., Trokovic, R., Sharma, V., Ylikallio, E., & Tyynismaa, H. (2021). Threshold of heteroplasmic truncating MT-ATP6 mutation in reprogramming, Notch hyperactivation and motor neuron metabolism. *Human Molecular Genetics*. https://doi.org/10.1093/hmg/ddab299

Kim, J. Y., Woo, S. Y., Hong, Y. Bin, Choi, H., Kim, J., Choi, H., Mook-Jung, I., Ha, N., Kyung, J., Koo, S. K., Jung, S. C., & Choi, B. O. (2016). HDAC6 Inhibitors Rescued the Defective Axonal Mitochondrial Movement in Motor Neurons Derived from the Induced Pluripotent Stem Cells of Peripheral Neuropathy Patients with HSPB1 Mutation. *Stem Cells International*, *2016*. https://doi.org/10.1155/2016/9475981

Kim, S. K., Kim, H., Yang, Y. R., Suh, P. G., & Chang, J. S. (2011). Phosphatidylinositol phosphates directly bind to neurofilament light chain (NF-L) for the regulation of NF-L self assembly. *Experimental and Molecular Medicine*, *43*(3), 153–160. https://doi.org/10.3858/emm.2011.43.3.019

Kim, Y. G., Cha, J., & Chandrasegaran, S. (1996). Hybrid restriction enzymes: Zinc finger fusions to Fok I cleavage domain. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(3), 1156–1160. https://doi.org/10.1073/pnas.93.3.1156

Kirkcaldie, M. T. K., & Dwyer, S. T. (2017). The third wave: Intermediate filaments in the maturing nervous system. In *Molecular and Cellular Neuroscience* (Vol. 84, pp.

68–76). Academic Press Inc. https://doi.org/10.1016/j.mcn.2017.05.010

Klein, C. J. (2020). Charcot-Marie-Tooth Disease and Other Hereditary Neuropathies. *Continuum (Minneapolis, Minn.)*, *26*(5), 1224–1256. https://doi.org/10.1212/CON.0000000000000927

Klein, C. J., Wu, Y., Vogel, P., Goebel, H. H., Bönnemann, C., Zukosky, K., Botuyan, M. V., Duan, X., Middha, S., Atkinson, E. J., Mer, G., & Dyck, P. J. (2014). Ubiquitin ligase defect by DCAF8 mutation causes HMSN2 with giant axons. *Neurology*, *82*(10), 873–878. https://doi.org/10.1212/WNL.0000000000000206

Klim, J. R., Williams, L. A., Limone, F., Guerra San Juan, I., Davis-Dusenbery, B. N., Mordes, D. A., Burberry, A., Steinbaugh, M. J., Gamage, K. K., Kirchner, R., Moccia, R., Cassel, S. H., Chen, K., Wainger, B. J., Woolf, C. J., & Eggan, K. (2019). ALS-implicated protein TDP-43 sustains levels of STMN2, a mediator of motor neuron growth and repair. *Nature Neuroscience*, *22*(2), 167–179. https://doi.org/10.1038/s41593-018-0300-4

Koch, T., Schultz, P., Williams, R., & Lampert, P. (1977). Giant axonal neuropathy: A childhood disorder of microfilaments. *Annals of Neurology*, *1*(5), 438–451. https://doi.org/10.1002/ana.410010507

Kornak, U., Mademan, I., Schinke, M., Voigt, M., Krawitz, P., Hecht, J., Barvencik, F., Schinke, T., Giesselmann, S., Beil, F. T., Pou-Serradell, A., Vilchez, J. J., Beetz, C., Deconinck, T., Timmerman, V., Kaether, C., De Jonghe, P., Hubner, C. A., Gal, A., … Kurth, I. (2014). Sensory neuropathy with bone destruction due to a mutation in the membrane-shaping atlastin GTPase 3. *Brain : A Journal of Neurology*, *137*(Pt 3), 683–692. https://doi.org/10.1093/brain/awt357 [doi]

Krichevsky, A. M., & Kosik, K. S. (2001). Neuronal RNA Granules: A Link between RNA Localization and Stimulation-Dependent Translation. *Neuron*, *32*(4), 683–696. https://doi.org/10.1016/S0896-6273(01)00508-6

Križ, J., Zhu, Q., Julien, J. P., & Padjen, A. L. (2000). Electrophysiological properties of axons in mice lacking neurofilament subunit genes: Disparity between conduction velocity and axon diameter in absence of NF-H. *Brain Research*, *885*(1), 32–44. https://doi.org/10.1016/S0006-8993(00)02899-7

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., Fitzhugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., Levine, R., McEwan, P., … Morgan, M. J. (2001). Initial sequencing and analysis of the human genome. *Nature*, *409*(6822), 860–921. https://doi.org/10.1038/35057062

laurent, G., Carlier, M. B., Rollman, B., Van Hoof, F., & Tulkens, P. (1982). Mechanism of aminoglycoside-induced lysosomal phospholipidosis: In vitro and in vivo studies with Gentamicin and Amikacin. *Biochemical Pharmacology*, *31*(23), 3861–3870. https://doi.org/10.1016/0006-2952(82)90303-3

Lee, S., Lee, B., Lee, J. W., & Lee, S. K. (2009). Retinoid Signaling and Neurogenin2 Function Are Coupled for the Specification of Spinal Motor Neurons through a

Chromatin Modifier CBP. *Neuron*, *62*(5), 641–654. https://doi.org/10.1016/j.neuron.2009.04.025

Leterrier, C., Dubey, P., & Roy, S. (2017). The nano-architecture of the axonal cytoskeleton. *Nature Reviews Neuroscience 2017 18:12*, *18*(12), 713–726. https://doi.org/10.1038/nrn.2017.129

Li, J., Parker, B., Martyn, C., Natarajan, C., & Guo, J. (2013). The PMP22 gene and its related diseases. In *Molecular neurobiology* (Vol. 47, Issue 2, pp. 673–698). Springer. https://doi.org/10.1007/s12035-012-8370-x

Li, X. J., Du, Z. W., Zarnowska, E. D., Pankratz, M., Hansen, L. O., Pearce, R. A., & Zhang, S. C. (2005). Specification of motoneurons from human embryonic stem cells. *Nature Biotechnology*, *23*(2), 215–221. https://doi.org/10.1038/nbt1063

Liewluck, T., & Milone, M. (2018). Untangling the complexity of limb-girdle muscular dystrophies. *Muscle and Nerve*, *58*(2), 167–177. https://doi.org/10.1002/mus.26077

Liu, J. P., Laufer, E., & Jessell, T. M. (2001). Assigning the positional identity of spinal motor neurons: Rostrocaudal patterning of Hox-c expression by FGFs, Gdf11, and retinoids. *Neuron*, *32*(6), 997–1012. https://doi.org/10.1016/S0896-6273(01)00544-X

Liu, X., Xu, Y., An, M., & Zeng, Q. (2019). The risk factors for diabetic peripheral neuropathy: A meta-analysis. *PLoS ONE*, *14*(2). https://doi.org/10.1371/journal.pone.0212574

Lornage, X., Schartner, V., Balbueno, I., Biancalana, V., Willis, T., Echaniz-Laguna, A., Scheidecker, S., Quinlivan, R., Fardeau, M., Malfatti, E., Lannes, B., Sewry, C., Romero, N. B., Laporte, J., & Böhm, J. (2019). Clinical, histological, and genetic characterization of PYROXD1-related myopathy. *Acta Neuropathologica Communications*, *7*(1), 138. https://doi.org/10.1186/s40478-019-0781-8

Maciel, R., Correa, R., Bosso Taniguchi, J., Prufer Araujo, I., & Saporta, M. A. (2020). Human Tridimensional Neuronal Cultures for Phenotypic Drug Screening in Inherited Peripheral Neuropathies. *Clinical Pharmacology & Therapeutics*, *107*(5), 1231–1239. https://doi.org/10.1002/CPT.1718

Malik, V., Rodino Klapac, L. R., Viollet, L., & Mendell, J. R. (2010). Aminoglycoside-induced mutation suppression (stop codon readthrough) as a therapeutic strategy for Duchenne muscular dystrophy. In *Therapeutic Advances in Neurological Disorders* (Vol. 3, Issue 6, pp. 379–389). Ther Adv Neurol Disord. https://doi.org/10.1177/1756285610388693

Manganelli, F., Parisi, S., Nolano, M., Miceli, F., Tozza, S., Pisciotta, C., Iodice, R., Provitera, V., Cicatiello, R., Zuchner, S., Taglialatela, M., Russo, T., & Santoro, L. (2019). Insights into the pathogenesis of ATP1A1-related CMT disease using patient-specific iPSCs. *Journal of the Peripheral Nervous System*, *24*(4), 330–339. https://doi.org/10.1111/jns.12357

Markworth, R., Bähr, M., & Burk, K. (2021). Held Up in Traffic—Defects in the Trafficking Machinery in Charcot-Marie-Tooth Disease. *Frontiers in Molecular Neuroscience*, *14*, 167. https://doi.org/10.3389/fnmol.2021.695294

Martin, G. R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proceedings of the National Academy of Sciences of the United States of America*, *78*(12 II), 7634–7638. https://doi.org/10.1073/pnas.78.12.7634

Martins-Dias, P., & Romão, L. (2021). Nonsense suppression therapies in human genetic diseases. In *Cellular and Molecular Life Sciences* (Vol. 78, Issue 10, pp. 4677–4701). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/s00018-021-03809-7

Marttila, M., Kytövuori, L., Helisalmi, S., Kallio, M., Laitinen, M., Hiltunen, M., Kärppä, M., & Majamaa, K. (2017). Molecular Epidemiology of Charcot-Marie-Tooth Disease in Northern Ost. *Neuroepidemiology*, *49*(1–2), 34–39. https://doi.org/10.1159/000478860

Maury, Y., Côme, J., Piskorowski, R. A., Salah-Mohellibi, N., Chevaleyre, V., Peschanski, M., Martinat, C., & Nedelec, S. (2015). Combinatorial analysis of developmental cues efficiently converts human pluripotent stem cells into multiple neuronal subtypes. *Nature Biotechnology*, *33*(1), 89–96. https://doi.org/10.1038/nbt.3049

Meyer, N., Richter, N., Fan, Z., Siemonsmeier, G., Pivneva, T., Jordan, P., Steinhäuser, C., Semtner, M., Nolte, C., & Kettenmann, H. (2018). Oligodendrocytes in the Mouse Corpus Callosum Maintain Axonal Function by Delivery of Glucose. *Cell Reports*, *22*(9), 2383–2394. https://doi.org/10.1016/j.celrep.2018.02.022

Michael, E., Sofou, K., Wahlgren, L., Kroksmark, A. K., & Tulinius, M. (2021). Long term treatment with ataluren—the Swedish experience. *BMC Musculoskeletal Disorders*, *22*(1). https://doi.org/10.1186/s12891-021-04700-z

Millecamps, S., Gowing, G., Corti, O., Mallet, J., & Julien, J.-P. (2007). Conditional NF-L Transgene Expression in Mice for In Vivo Analysis of Turnover and Transport Rate of Neurofilaments. *The Journal of Neuroscience*, *27*(18), 4947. https://doi.org/10.1523/JNEUROSCI.5299-06.2007

Miressi, F., Benslimane, N., Favreau, F., Rassat, M., Richard, L., Bourthoumieu, S., Laroche, C., Magy, L., Magdelaine, C., Sturtz, F., Lia, A. S., & Faye, P. A. (2021). Gdap1 involvement in mitochondrial function and oxidative stress, investigated in a charcot-marie-tooth model of hipscs-derived motor neurons. *Biomedicines*, *9*(8). https://doi.org/10.3390/biomedicines9080945

Mizuguchi, R., Sugimori, M., Takebayashi, H., Kosako, H., Nagao, M., Yoshida, S., Nabeshima, Y. I., Shimamura, K., & Nakafuku, M. (2001). Combinatorial roles of Olig2 and Neurogenin2 in the coordinated induction of pan-neuronal and subtype-specific properties of motoneurons. *Neuron*, *31*(5), 757–771. https://doi.org/10.1016/S0896-6273(01)00413-5

Morrison, O., & Thakur, J. (2021). Molecular complexes at euchromatin, heterochromatin and centromeric chromatin. *International Journal of Molecular Sciences*, *22*(13). https://doi.org/10.3390/ijms22136922

Moscou, M. J., & Bogdanove, A. J. (2009). A simple cipher governs DNA recognition by TAL effectors. *Science*, *326*(5959), 1501. https://doi.org/10.1126/science.1178817

Mou, Y., Dein, J., Chen, Z., Jagdale, M., & Li, X. J. (2021). MFN2 Deficiency Impairs Mitochondrial Transport and Downregulates Motor Protein Expression in Human Spinal Motor Neurons. *Frontiers in Molecular Neuroscience*, *14*. https://doi.org/10.3389/fnmol.2021.727552

Murphy, S. M., Laura, M., Fawcett, K., Pandraud, A., Liu, Y. T., Davidson, G. L., Rossor, A. M., Polke, J. M., Castleman, V., Manji, H., Lunn, M. P. T., Bull, K., Ramdharry, G., Davis, M., Blake, J. C., Houlden, H., & Reilly, M. M. (2012). Charcot-Marie-Tooth disease: Frequency of genetic subtypes and guidelines for genetic testing. *Journal of Neurology, Neurosurgery and Psychiatry*, *83*(7), 706–710. https://doi.org/10.1136/jnnp-2012-302451

Nagappa, M., Bindu, P. S., Sinha, S., Mathuranath, P. S., & Taly, A. B. (2018). Exome sequencing in adult neurology practice: Challenges and rewards in a mixed resource setting. *Clinical Neurology and Neurosurgery*, *174*, 48–56. https://doi.org/10.1016/j.clineuro.2018.09.012

Nagel-Wolfrum, K., Möller, F., Penner, I., Baasov, T., & Wolfrum, U. (2016). Targeting Nonsense Mutations in Diseases with Translational Read-Through-Inducing Drugs (TRIDs). *BioDrugs*, *30*(2), 49–74. https://doi.org/10.1007/s40259-016-0157-6

Nam, D. E., Jung, S.-C., Yoo, D. H., Choi, S. S., Seo, S.-Y., Kim, G. H., Kim, S. J., Nam, S. H., Choi, B.-O., & Chung, K. W. (2017). Axonal Charcot-Marie-Tooth neuropathy concurrent with distal and proximal weakness by translational elongation of the 3′ UTR in NEFH. *Journal of the Peripheral Nervous System*, *22*(3), 200–207. https://doi.org/10.1111/JNS.12223

Nave, K. A. (2010). Myelination and the trophic support of long axons. In *Nature Reviews Neuroscience* (Vol. 11, Issue 4, pp. 275–283). Nature Publishing Group. https://doi.org/10.1038/nrn2797

Nielsen, R., Akey, J. M., Jakobsson, M., Pritchard, J. K., Tishkoff, S., & Willerslev, E. (2017). Tracing the peopling of the world through genomics. *Nature*, *541*(7637), 302. https://doi.org/10.1038/NATURE21347

Nixon, R. A., Paskevich, P. A., Sihag, R. K., & Thayer, C. Y. (1994). Phosphorylation on carboxyl terminus domains of neurofilament proteins in retinal ganglion cell neurons in vivo: Influences on regional neurofilament accumulation, interneurofilament spacing, and axon caliber. *Journal of Cell Biology*, *126*(4), 1031–1046. https://doi.org/10.1083/jcb.126.4.1031

Nixon, Ralph A., & Shea, T. B. (1992). Dynamics of neuronal intermediate filaments: A developmental perspective. *Cell Motility and the Cytoskeleton*, *22*(2), 81–91. https://doi.org/10.1002/cm.970220202

Nixon, Ralph A., & Sihag, R. K. (1991). Neurofilament phosphorylation: a new look at regulation and function. In *Trends in Neurosciences* (Vol. 14, Issue 11, pp. 501–506). Elsevier Current Trends. https://doi.org/10.1016/0166-2236(91)90062-Y

Norio, R. (2003a). Finnish Disease Heritage II: Population prehistory and genetic roots of Finns. In *Human Genetics* (Vol. 112, Issues 5–6, pp. 457–469). Springer Verlag. https://doi.org/10.1007/s00439-002-0876-2

Norio, R. (2003b). The Finnish disease heritage III: The individual diseases. In *Human Genetics* (Vol. 112, Issues 5–6, pp. 470–526). Springer Verlag. https://doi.org/10.1007/s00439-002-0877-1

Norwood, F. L. M., Harling, C., Chinnery, P. F., Eagle, M., Bushby, K., & Straub, V. (2009). Prevalence of genetic muscle disease in Northern England: In-depth analysis of a muscle clinic population. *Brain*, *132*(11), 3175–3186. https://doi.org/10.1093/brain/awp236

O'Grady, G. L., Best, H. A., Sztal, T. E., Schartner, V., Sanjuan-Vazquez, M., Donkervoort, S., Abath Neto, O., Sutton, R. B., Ilkovski, B., Romero, N. B., Stojkovic, T., Dastgir, J., Waddell, L. B., Boland, A., Hu, Y., Williams, C., Ruparelia, A. A., Maisonobe, T., Peduto, A. J., … Cooper, S. T. (2016). Variants in the Oxidoreductase PYROXD1 Cause Early-Onset Myopathy with Internalized Nuclei and Myofibrillar Disorganization. *American Journal of Human Genetics*, *99*(5), 1086–1105. https://doi.org/S0002-9297(16)30385-8 [pii]

Ohara, O., Gahara, Y., Miyake, T., Teraoka, H., & Kitamura, T. (1993). Neurofilament deficiency in quail caused by nonsense mutation in neurofilament-L gene. *Journal of Cell Biology*, *121*(2), 387–395. https://doi.org/10.1083/jcb.121.2.387

Ohara, R., Imamura, K., Morii, F., Egawa, N., Tsukita, K., Enami, T., Shibukawa, R., Mizuno, T., Nakagawa, M., & Inoue, H. (2017). Modeling Drug-Induced Neuropathy Using Human iPSCs for Predictive Toxicology. *Clinical Pharmacology and Therapeutics*, *101*(6), 754–762. https://doi.org/10.1002/cpt.562

Ozawa, E., Mizuno, Y., Hagiwara, Y., Sasaoka, T., & Yoshida, M. (2005). Molecular and cell biology of the sarcoglycan complex. *Muscle and Nerve*, *32*(5), 563–576. https://doi.org/10.1002/mus.20349

Özyilmaz, B., Kirbiyik, Ö., Özdemir, T. R., Kaya Özer, Ö., Kutbay, Y. B., Erdogan, K. M., Güvenç, M. S., Kale, M. Y., Gazeteci, H., Kiliç, B., Sertpoyraz, F., Diniz, G., Baydan, F., Gençpinar, P., Dündar, N. O., & Yiş, U. (2019). Impact of next-generation sequencing panels in the evaluation of limb-girdle muscular dystrophies. *Annals of Human Genetics*, *83*(5), 331–347. https://doi.org/10.1111/ahg.12319

Pachter, J. S., & Liem, R. K. H. (1984). The differential appearance of neurofilament triplet polypeptides in the developing rat optic nerve. *Developmental Biology*, *103*(1), 200–210. https://doi.org/10.1016/0012-1606(84)90021-6

Parodi, L., Coarelli, G., Stevanin, G., Brice, A., & Durr, A. (2018). Hereditary ataxias and paraparesias: clinical and genetic update. *Current Opinion in Neurology*, *31*(4), 462–471. https://doi.org/10.1097/WCO.0000000000000585 [doi]

Pastinen, T., Perola, M., Ignatius, J., Sabatti, C., Tainola, P., Levander, M., Syvänen, A. C., & Peltonen, L. (2001). Dissecting a population genome for targeted screening of disease mutations. *Human Molecular Genetics*, *10*(26), 2961–2972.

https://doi.org/10.1093/hmg/10.26.2961

Penttilä, S., Jokela, M., Bouquin, H., Saukkonen, A. M., Toivanen, J., & Udd, B. (2015). Late onset spinal motor neuronopathy is caused by mutation in CHCHD10. *Annals of Neurology*, *77*(1), 163–172. https://doi.org/10.1002/ana.24319

Penttila, S., Palmio, J., Suominen, T., Raheem, O., Evila, A., Muelas Gomez, N., Tasca, G., Waddell, L. B., Clarke, N. F., Barboi, A., Hackman, P., & Udd, B. (2012). Eight new mutations and the expanding phenotype variability in muscular dystrophy caused by ANO5. *Neurology*, *78*(12), 897–903. https://doi.org/10.1212/WNL.0b013e31824c4682 [doi]

Perez-Siles, G., Cutrupi, A., Ellis, M., Screnci, R., Mao, D., Uesugi, M., Yiu, E. M., Ryan, M. M., Choi, B. O., Nicholson, G., & Kennerson, M. L. (2020). Energy metabolism and mitochondrial defects in X-linked Charcot-Marie-Tooth (CMTX6) iPSC-derived motor neurons with the p.R158H PDK3 mutation. *Scientific Reports*, *10*(1). https://doi.org/10.1038/s41598-020-66266-5

Perez-Siles, Gonzalo, Cutrupi, A., Ellis, M., Kuriakose, J., La Fontaine, S., Mao, D., Uesugi, M., Takata, R. I., Speck-Martins, C. E., Nicholson, G., & Kennerson, M. L. (2020). Modelling the pathogenesis of X-linked distal hereditary motor neuropathy using patient-derived iPSCs. *DMM Disease Models and Mechanisms*, *13*(2). https://doi.org/10.1242/dmm.041541

Perlman, R. L. (2016). Mouse Models of Human Disease: An Evolutionary Perspective. *Evolution, Medicine, and Public Health*, *2016*(1), eow014. https://doi.org/10.1093/emph/eow014

Perrot, R., & Julien, J.-P. (2009). Real-time imaging reveals defects of fast axonal transport induced by disorganization of intermediate filaments. *The FASEB Journal*, *23*(9), 3213–3225. https://doi.org/10.1096/fj.09-129585

Perry, G., Stewart, D., Friedman, R., Manetto, V., Autilio-Gambetti, L., & Gambetti, P. (1987). Filaments of Pick's bodies contain altered cytoskeletal elements. *American Journal of Pathology*, *127*(3), 559–568. /pmc/articles/PMC1899769/?report=abstract

Picci, C., Wong, V. S. C., Costa, C. J., McKinnon, M. C., Goldberg, D. C., Swift, M., Alam, N. M., Prusky, G. T., Shen, S., Kozikowski, A. P., Willis, D. E., & Langley, B. (2020). HDAC6 inhibition promotes α-tubulin acetylation and ameliorates CMT2A peripheral neuropathy in mice. *Experimental Neurology*, *328*. https://doi.org/10.1016/j.expneurol.2020.113281

Pipis, M., Cortese, A., Polke, J. M., Poh, R., Vandrovcova, J., Laura, M., Skorupinska, M., Jacquier, A., Juntas-Morales, R., Latour, P., Petiot, P., Sole, G., Fromes, Y., Shah, S., Blake, J., Choi, B.-O., Chung, K. W., Stojkovic, T., Rossor, A. M., & Reilly, M. M. (2022). Charcot-Marie-Tooth disease type 2CC due to NEFH variants causes a progressive, non-length-dependent, motor-predominant phenotype . *Journal of Neurology, Neurosurgery & Psychiatry*, *93*(1), 48–56. https://doi.org/10.1136/jnnp-2021-327186

Pipis, M., Rossor, A. M., Laura, M., & Reilly, M. M. (2019). Next-generation sequencing

in Charcot–Marie–Tooth disease: opportunities and challenges. In *Nature Reviews Neurology* (Vol. 15, Issue 11, pp. 644–656). Nature Publishing Group. https://doi.org/10.1038/s41582-019-0254-5

Posey, J. E., Harel, T., Liu, P., Rosenfeld, J. A., James, R. A., Coban Akdemir, Z. H., Walkiewicz, M., Bi, W., Xiao, R., Ding, Y., Xia, F., Beaudet, A. L., Muzny, D. M., Gibbs, R. A., Boerwinkle, E., Eng, C. M., Sutton, V. R., Shaw, C. A., Plon, S. E., … Lupski, J. R. (2017). Resolution of Disease Phenotypes Resulting from Multilocus Genomic Variation. *New England Journal of Medicine*, *376*(1), 21–31. https://doi.org/10.1056/nejmoa1516767

Posey, J. E., Rosenfeld, J. A., James, R. A., Bainbridge, M., Niu, Z., Wang, X., Dhar, S., Wiszniewski, W., Akdemir, Z. H. C., Gambin, T., Xia, F., Person, R. E., Walkiewicz, M., Shaw, C. A., Sutton, V. R., Beaudet, A. L., Muzny, D., Eng, C. M., Yang, Y., … Plon, S. E. (2016). Molecular diagnostic experience of whole-exome sequencing in adult patients. *Genetics in Medicine*, *18*(7), 678–685. https://doi.org/10.1038/gim.2015.142

Rao, M. V., Engle, L. J., Mohan, P. S., Yuan, A., Qiu, D., Cataldo, A., Hassinger, L., Jacobsen, S., Lee, V. M. Y., Andreadis, A., Julien, J. P., Bridgman, P. C., & Nixon, R. A. (2002). Myosin Va binding to neurofilaments is essential for correct myosin Va distribution and transport and neurofilament density. *Journal of Cell Biology*, *159*(2), 279–289. https://doi.org/10.1083/jcb.200205062

Rao, M. V., Mohan, P. S., Kumar, A., Yuan, A., Montagna, L., Campbell, J., Veeranna, Espreafico, E. M., Julien, J. P., & Nixon, R. A. (2011). The myosin Va head domain binds to the neurofilament-L rod and modulates endoplasmic reticulum (ER) content and distribution within axons. *PLoS ONE*, *6*(2). https://doi.org/10.1371/journal.pone.0017087

Rao, M. V., Yuan, A., Campbell, J., Kumar, A., & Nixon, R. A. (2012). The C-Terminal Domains of NF-H and NF-M Subunits Maintain Axonal Neurofilament Content by Blocking Turnover of the Stationary Neurofilament Network. *PLoS ONE*, *7*(9). https://doi.org/10.1371/journal.pone.0044320

Rebelo, A. P., Abrams, A. J., Cottenie, E., Horga, A., Gonzalez, M., Bis, D. M., Sanchez-Mejias, A., Pinto, M., Buglo, E., Markel, K., Prince, J., Laura, M., Houlden, H., Blake, J., Woodward, C., Sweeney, M. G., Holton, J. L., Hanna, M., Dallman, J. E., … Zuchner, S. (2016). Cryptic Amyloidogenic Elements in the 3′ UTRs of Neurofilament Genes Trigger Axonal Neuropathy. *American Journal of Human Genetics*, *98*(4), 597–614. https://doi.org/10.1016/j.ajhg.2016.02.022

Reddy, H. M., Cho, K. A., Lek, M., Estrella, E., Valkanas, E., Jones, M. D., Mitsuhashi, S., Darras, B. T., Amato, A. A., Lidov, H. G., Brownstein, C. A., Margulies, D. M., Yu, T. W., Salih, M. A., Kunkel, L. M., MacArthur, D. G., & Kang, P. B. (2017). The sensitivity of exome sequencing in identifying pathogenic mutations for LGMD in the United States. *Journal of Human Genetics*, *62*(2), 243–252. https://doi.org/10.1038/jhg.2016.116 [doi]

Reed, C. B., Feltri, M. L., & Wilson, E. R. (2021). Peripheral glia diversity. *Journal of*

Anatomy, joa.13484. https://doi.org/10.1111/joa.13484

Reid Alderson, T., Adriaenssens, E., Asselbergh, B., Pritišanac, I., Van Lent, J., Gastall, H. Y., Wälti, M. A., Louis, J. M., Timmerman, V., Baldwin, A. J., & LP Benesch, J. (2021). A weakened interface in the P182L variant of HSP27 associated with severe Charcot-Marie-Tooth neuropathy causes aberrant binding to interacting proteins. *The EMBO Journal*, *40*(8). https://doi.org/10.15252/embj.2019103811

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W. W., Hegde, M., Lyon, E., Spector, E., Voelkerding, K., Rehm, H. L., & Committee, A. L. Q. A. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in Medicine : Official Journal of the American College of Medical Genetics*, *17*(5), 405–424. https://doi.org/10.1038/gim.2015.30 [doi]

Richmond, A., & Yingjun, S. (2008). Mouse xenograft models vs GEM models for human cancer therapeutics. *DMM Disease Models and Mechanisms*, *1*(2–3), 78–82. https://doi.org/10.1242/dmm.000976

Rönkkö, J., Molchanova, S., Revah-Politi, A., Pereira, E. M., Auranen, M., Toppila, J., Kvist, J., Ludwig, A., Neumann, J., Bultynck, G., Humblet-Baron, S., Liston, A., Paetau, A., Rivera, C., Harms, M. B., Tyynismaa, H., & Ylikallio, E. (2020). Dominant mutations in ITPR3 cause Charcot-Marie-Tooth disease. *Annals of Clinical and Translational Neurology*, *7*(10), 1962–1972. https://doi.org/10.1002/acn3.51190

Roos, W. P., Thomas, A. D., & Kaina, B. (2016). DNA damage and the balance between survival and death in cancer biology. *Nature Reviews Cancer*, *16*(1), 20–33. https://doi.org/10.1038/nrc.2015.2

Rossor, A. M., Polke, J. M., Houlden, H., & Reilly, M. M. (2013). Clinical implications of genetic advances in Charcot-Marie-Tooth disease. *Nature Reviews.Neurology*, *9*(10), 562–571. https://doi.org/10.1038/nrneurol.2013.179 [doi]

Rossor, A. M., Tomaselli, P. J., & Reilly, M. M. (2016). Recent advances in the genetic neuropathies. *Current Opinion in Neurology*, *29*(5), 537–548. https://doi.org/10.1097/WCO.0000000000000373 [doi]

Rouet, P., Smih, F., & Jasin, M. (1994). Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Molecular and Cellular Biology*, *14*(12), 8096–8106. https://doi.org/10.1128/mcb.14.12.8096-8106.1994

Saha, M., Reddy, H. M., Salih, M. A., Estrella, E., Jones, M. D., Mitsuhashi, S., Cho, K. A., Suzuki-Hatano, S., Rizzo, S. A., Hamad, M. H., Mukhtar, M. M., Hamed, A. A., Elseed, M. A., Lek, M., Valkanas, E., Macarthur, D. G., Kunkel, L. M., Pacak, C. A., Draper, I., & Kang, P. B. (2018). Impact of PYROXD1 deficiency on cellular respiration and correlations with genetic analyses of limb-girdle muscular dystrophy in Saudi Arabia and Sudan. *Physiological Genomics*, *50*(11), 929–939. https://doi.org/10.1152/physiolgenomics.00036.2018

Sainio, M. T., Välipakka, S., Rinaldi, B., Lapatto, H., Paetau, A., Ojanen, S., Brilhante, V., Jokela, M., Huovinen, S., Auranen, M., Palmio, J., Friant, S., Ylikallio, E., Udd, B., & Tyynismaa, H. (2019). Recessive PYROXD1 mutations cause adult-onset limb-girdle-type muscular dystrophy. *Journal of Neurology*, *266*(2). https://doi.org/10.1007/s00415-018-9137-8

Sainio, M. T., Ylikallio, E., Mäenpää, L., Lahtela, J., Mattila, P., Auranen, M., Palmio, J., & Tyynismaa, H. (2018). Absence of NEFL in patient-specific neurons in early-onset Charcot-Marie-Tooth neuropathy. *Neurology: Genetics*, *4*(3). https://doi.org/10.1212/NXG.0000000000000244

Salmela, E., Lappalainen, T., Fransson, I., Andersen, P. M., Dahlman-Wright, K., Fiebig, A., Sistonen, P., Savontaus, M. L., Schreiber, S., Kere, J., & Lahermo, P. (2008). Genome-Wide Analysis of Single Nucleotide Polymorphisms Uncovers Population Structure in Northern Europe. *PLoS ONE*, *3*(10). https://doi.org/10.1371/JOURNAL.PONE.0003519

Salzer, J. L. (2015). Schwann cell myelination. *Cold Spring Harbor Perspectives in Biology*, *7*(8). https://doi.org/10.1101/cshperspect.a020529

Sances, S., Bruijn, L. I., Chandran, S., Eggan, K., Ho, R., Klim, J. R., Livesey, M. R., Lowry, E., Macklis, J. D., Rushton, D., Sadegh, C., Sareen, D., Wichterle, H., Zhang, S. C., & Svendsen, C. N. (2016). Modeling ALS with motor neurons derived from human induced pluripotent stem cells. *Nature Neuroscience*, *19*(4), 542–553. https://doi.org/10.1038/nn.4273

Sanger, F., & Coulson, A. R. (1975). A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of Molecular Biology*, *94*(3). https://doi.org/10.1016/0022-2836(75)90213-2

Saporta, M. A., Dang, V., Volfson, D., Zou, B., Xie, X. S., Adebola, A., Liem, R. K., Shy, M., & Dimos, J. T. (2015). Axonal Charcot-Marie-Tooth disease patient-derived motor neurons demonstrate disease-specific phenotypes including abnormal electrophysiological properties. *Experimental Neurology*, *263*, 190–199. https://doi.org/10.1016/j.expneurol.2014.10.005

Sasaki, T., Gotow, T., Shiozaki, M., Sakaue, F., Saito, T., Julien, J. P., Uchiyama, Y., & Hisanaga, S. I. (2006). Aggregate formation and phosphorylation of neurofilament-L Pro22 Charcot-Marie-Tooth disease mutants. *Human Molecular Genetics*, *15*(6), 943–952. https://doi.org/10.1093/hmg/ddl011

Savarese, M., Sarparanta, J., Vihola, A., Udd, B., & Hackman, P. (2016). Increasing Role of Titin Mutations in Neuromuscular Disorders. *Journal of Neuromuscular Diseases*, *3*(3), 293–308. https://doi.org/10.3233/JND-160158

Schiavon, C. R., Shadel, G. S., & Manor, U. (2021). Impaired Mitochondrial Mobility in Charcot-Marie-Tooth Disease. *Frontiers in Cell and Developmental Biology*, *9*, 35. https://doi.org/10.3389/fcell.2021.624823

Schwarze, K., Buchanan, J., Taylor, J. C., & Wordsworth, S. (2018). Are whole-exome and whole-genome sequencing approaches cost-effective? A systematic review of the

literature. In *Genetics in Medicine* (Vol. 20, Issue 10, pp. 1122–1130). Nature Publishing Group. https://doi.org/10.1038/gim.2017.247

Senderek, J., Lassuthova, P., Kabzińska, D., Abreu, L., Baets, J., Beetz, C., Braathen, G. J., Brenner, D., Dalton, J., Dankwa, L., Deconinck, T., De Jonghe, P., Dräger, B., Eggermann, K., Ellis, M., Fischer, C., Stojkovic, T., Herrmann, D. N., Horvath, R., … Auer-Grumbach, M. (2020). The genetic landscape of axonal neuropathies in the middle-aged and elderly: Focus on MME. *Neurology*, *95*(24), e3163–e3179. https://doi.org/10.1212/WNL.0000000000011132

Sereda, M., Griffiths, I., Pühlhofer, A., Stewart, H., Rossner, M. J., Zimmermann, F., Magyar, J. P., Schneider, A., Hund, E., Meinck, H. M., Suter, U., & Nave, K. A. (1996). A transgenic rat model of Charcot-Marie-Tooth disease. *Neuron*, *16*(5), 1049–1060. https://doi.org/10.1016/S0896-6273(00)80128-2

Shah, N., Hou, Y. C. C., Yu, H. C., Sainger, R., Caskey, C. T., Venter, J. C., & Telenti, A. (2018). Identification of Misclassified ClinVar Variants via Disease Population Prevalence. *American Journal of Human Genetics*, *102*(4), 609–619. https://doi.org/10.1016/j.ajhg.2018.02.019

Shi, L., Huang, L., He, R., Huang, W., Wang, H., Lai, X., Zou, Z., Sun, J., Ke, Q., Zheng, M., Lu, X., Pei, Z., Su, H., Xiang, A. P., Li, W., & Yao, X. (2018). Modeling the Pathogenesis of Charcot-Marie-Tooth Disease Type 1A Using Patient-Specific iPSCs. *Stem Cell Reports*, *10*(1), 120–133. https://doi.org/10.1016/j.stemcr.2017.11.013

Shi, Y., Lin, S., Staats, K. A., Li, Y., Chang, W. H., Hung, S. T., Hendricks, E., Linares, G. R., Wang, Y., Son, E. Y., Wen, X., Kisler, K., Wilkinson, B., Menendez, L., Sugawara, T., Woolwine, P., Huang, M., Cowan, M. J., Ge, B., … Ichida, J. K. (2018). Haploinsufficiency leads to neurodegeneration in C9ORF72 ALS/FTD human induced motor neurons. *Nature Medicine*, *24*(3), 313–325. https://doi.org/10.1038/nm.4490

Shimojo, D., Onodera, K., Doi-Torii, Y., Ishihara, Y., Hattori, C., Miwa, Y., Tanaka, S., Okada, R., Ohyama, M., Shoji, M., Nakanishi, A., Doyu, M., Okano, H., & Okada, Y. (2015). Rapid, efficient, and simple motor neuron differentiation from human pluripotent stem cells. *Molecular Brain*, *8*(1). https://doi.org/10.1186/s13041-015-0172-4

Shirasaki, R., & Pfaff, S. L. (2002). Transcriptional codes and the control of neuronal identity. *Annual Review of Neuroscience*, *25*, 251–281. https://doi.org/10.1146/annurev.neuro.25.112701.142916

Shribman, S., Reid, E., Crosby, A. H., Houlden, H., & Warner, T. T. (2019). Hereditary spastic paraplegia: from diagnosis to emerging therapeutic approaches. *The Lancet Neurology*, *18*(12), 1136–1146. https://doi.org/10.1016/S1474-4422(19)30235-2

Sihag, R. K., Inagaki, M., Yamaguchi, T., Shea, T. B., & Pant, H. C. (2007). Role of phosphorylation on the structural dynamics and function of types III and IV intermediate filaments. In *Experimental Cell Research* (Vol. 313, Issue 10, pp. 2098–2109). Exp Cell Res. https://doi.org/10.1016/j.yexcr.2007.04.010

Skre, H. (1974). Genetic and clinical aspects of Charcot-Marie-Tooth's disease. *Clinical Genetics*, *6*(2), 98–118. https://doi.org/10.1111/j.1399-0004.1974.tb00638.x

Smith, G. M., & Gallo, G. (2018). The role of mitochondria in axon development and regeneration. In *Developmental Neurobiology* (Vol. 78, Issue 3, pp. 221–237). John Wiley and Sons Inc. https://doi.org/10.1002/dneu.22546

Son, E. Y., Ichida, J. K., Wainger, B. J., Toma, J. S., Rafuse, V. F., Woolf, C. J., & Eggan, K. (2011). Conversion of mouse and human fibroblasts into functional spinal motor neurons. *Cell Stem Cell*, *9*(3), 205–218. https://doi.org/10.1016/j.stem.2011.07.014

Splinter, K., Adams, D. R., Bacino, C. A., Bellen, H. J., Bernstein, J. A., Cheatle-Jarvela, A. M., Eng, C. M., Esteves, C., Gahl, W. A., Hamid, R., Jacob, H. J., Kikani, B., Koeller, D. M., Kohane, I. S., Lee, B. H., Loscalzo, J., Luo, X., McCray, A. T., Metz, T. O., … Ashley, E. A. (2018). Effect of Genetic Diagnosis on Patients with Previously Undiagnosed Disease. *New England Journal of Medicine*, *379*(22), 2131–2139. https://doi.org/10.1056/nejmoa1714458

Stark, Z., Schofield, D., Martyn, M., Rynehart, L., Shrestha, R., Alam, K., Lunke, S., Tan, T. Y., Gaff, C. L., & White, S. M. (2019). Does genomic sequencing early in the diagnostic trajectory make a difference? A follow-up study of clinical outcomes and cost-effectiveness. *Genetics in Medicine*, *21*(1), 173–180. https://doi.org/10.1038/s41436-018-0006-8

Stern, C. D. (2005). Neural induction: old problem, new findings, yet more questions. *Development*, *132*(9), 2007–2021. https://doi.org/10.1242/DEV.01794

Stifani, N. (2014). Motor neurons and the generation of spinal motor neuron diversity. In *Frontiers in Cellular Neuroscience* (Vol. 8, Issue OCT). Frontiers Media S.A. https://doi.org/10.3389/fncel.2014.00293

Stone, E. J., Kolb, S. J., & Brown, A. (2021). A review and analysis of the clinical literature on Charcot–Marie–Tooth disease caused by mutations in neurofilament protein L. *Cytoskeleton*, *78*(3), 97–110. https://doi.org/10.1002/cm.21676

Straub, V., Murphy, A., & Udd, B. (2018). 229th ENMC international workshop: Limb girdle muscular dystrophies – Nomenclature and reformed classification Naarden, the Netherlands, 17–19 March 2017. *Neuromuscular Disorders*, *28*(8), 702–710. https://doi.org/10.1016/j.nmd.2018.05.007

Styers, M. L., Salazar, G., Love, R., Peden, A. A., Kowalczyk, A. P., & Faundez, V. (2004). The endo-lysosomal sorting machinery interacts with the intermediate filament cytoskeleton. *Molecular Biology of the Cell*, *15*(12), 5369–5382. https://doi.org/10.1091/mbc.E04-03-0272

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., & Yamanaka, S. (2007). Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell*, *131*(5), 861–872. https://doi.org/10.1016/j.cell.2007.11.019

Takahashi, K., & Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell*, *126*(4), 663–676.

https://doi.org/10.1016/j.cell.2006.07.024

Takahashi, K., & Yamanaka, S. (2015). A developmental framework for induced pluripotency. *Development (Cambridge)*, *142*(19), 3274–3285. https://doi.org/10.1242/dev.114249

Takahashi, K., & Yamanaka, S. (2016). A decade of transcription factor-mediated reprogramming to pluripotency. *Nature Reviews Molecular Cell Biology*, *17*(3), 183–193. https://doi.org/10.1038/nrm.2016.8

Tang, B. S., Zhao, G. hua, Luo, W., Xia, K., Cai, F., Pan, Q., Zhang, R. xu, Zhang, F. feng, Liu, X. min, Chen, B., Zhang, C., Shen, L., Jiang, H., Long, Z. gao, & Dai, H. ping. (2005). Small heat-shock protein 22 mutated in autosomal dominant Charcot-Marie-Tooth disease type 2L. *Human Genetics*, *116*(3), 222–224. https://doi.org/10.1007/s00439-004-1218-3

Tarakci, H., & Berger, J. (2016). The sarcoglycan complex in skeletal muscle. In *Frontiers in Bioscience - Landmark* (Vol. 21, Issue 4, pp. 744–756). Frontiers in Bioscience. https://doi.org/10.2741/4418

Thomson, J. A. (1998). Embryonic stem cell lines derived from human blastocysts. *Science*, *282*(5391), 1145–1147. https://doi.org/10.1126/science.282.5391.1145

Torres-Ruiz, R., & Rodriguez-Perales, S. (2017). CRISPR-Cas9 technology: Applications and human disease modelling. *Briefings in Functional Genomics*, *16*(1), 4–12. https://doi.org/10.1093/bfgp/elw025

Trujillano, D., Bertoli-Avella, A. M., Kumar Kandaswamy, K., Weiss, M. E., Köster, J., Marais, A., Paknia, O., Schröder, R., Garcia-Aznar, J. M., Werber, M., Brandau, O., Calvo Del Castillo, M., Baldi, C., Wessel, K., Kishore, S., Nahavandi, N., Eyaid, W., Al Rifai, M. T., Al-Rumayyan, A., … Abou Jamra, R. (2017). Clinical exome sequencing: Results from 2819 samples reflecting 1000 families. *European Journal of Human Genetics*, *25*(2), 176–182. https://doi.org/10.1038/ejhg.2016.146

Tsang, Y. M., Chiong, F., Kuznetsov, D., Kasarskis, E., & Geula, C. (2000). Motor neurons are rich in non-phosphorylated neurofilaments: cross-species comparison and alterations in ALS. *Brain Research*, *861*(1), 45–58. https://doi.org/10.1016/S0006-8993(00)01954-5

Tuladhar, R., Yeu, Y., Tyler Piazza, J., Tan, Z., Rene Clemenceau, J., Wu, X., Barrett, Q., Herbert, J., Mathews, D. H., Kim, J., Hyun Hwang, T., & Lum, L. (2019). CRISPR-Cas9-based mutagenesis frequently provokes on-target mRNA misregulation. *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-019-12028-5

Uchida, A., Colakoglu, G., Wang, L., Monsma, P. C., & Brown, A. (2013). Severing and end-to-end annealing of neurofilaments in neurons. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(29), E2696. https://doi.org/10.1073/pnas.1221835110

Van Goethem, G., Luoma, P., Rantamäki, M., Al Memar, A., Kaakkola, S., Hackman, P., Krahe, R., Löfgren, A., Martin, J. J., De Jonghe, P., Suomalainen, A., Udd, B., & Van Broeckhoven, C. (2004). POLG mutations in neurodegenerative disorders with

ataxia but no muscle involvement. *Neurology*, *63*(7), 1251–1257. https://doi.org/10.1212/01.WNL.0000140494.58732.83

Van Lent, J., Verstraelen, P., Asselbergh, B., Adriaenssens, E., Mateiu, L., Verbist, C., De Winter, V., Eggermont, K., Van Den Bosch, L., De Vos, W. H., & Timmerman, V. (2021). Induced pluripotent stem cell-derived motor neurons of CMT type 2 patients reveal progressive mitochondrial dysfunction. *Brain*, *144*(8), 2471–2485. https://doi.org/10.1093/brain/awab226

Verhaart, I. E. C., Robertson, A., Wilson, I. J., Aartsma-Rus, A., Cameron, S., Jones, C. C., Cook, S. F., & Lochmüller, H. (2017). Prevalence, incidence and carrier frequency of 5q-linked spinal muscular atrophy - A literature review. *Orphanet Journal of Rare Diseases*, *12*(1). https://doi.org/10.1186/s13023-017-0671-8

Vierbuchen, T., Ostermeier, A., Pang, Z. P., Kokubu, Y., Südhof, T. C., & Wernig, M. (2010). Direct conversion of fibroblasts to functional neurons by defined factors. *Nature*, *463*(7284), 1035–1041. https://doi.org/10.1038/nature08797

Vierbuchen, T., & Wernig, M. (2011). Direct lineage conversions: Unnatural but useful? In *Nature Biotechnology* (Vol. 29, Issue 10, pp. 892–907). https://doi.org/10.1038/nbt.1946

Vissers, L. E. L. M., van Nimwegen, K. J. M., Schieving, J. H., Kamsteeg, E. J., Kleefstra, T., Yntema, H. G., Pfundt, R., van der Wilt, G. J., Krabbenborg, L., Brunner, H. G., van der Burg, S., Grutters, J., Veltman, J. A., & Willemsen, M. A. A. P. (2017). A clinical utility study of exome sequencing versus conventional genetic testing in pediatric neurology. *Genetics in Medicine : Official Journal of the American College of Medical Genetics*, *19*(9), 1055–1063. https://doi.org/10.1038/gim.2017.1 [doi]

Wagner, O. I., Lifshitz, J., Janmey, P. A., Linden, M., McIntosh, T. K., & Leterrier, J. F. (2003). Mechanisms of mitochondria-neurofilament interactions. *Journal of Neuroscience*, *23*(27), 9046–9058. https://doi.org/10.1523/jneurosci.23-27-09046.2003

Wang, Lei, Ho, C. L., Sun, D., Liem, R. K. H., & Brown, A. (2000). Rapid movement of axonal neurofilaments interrupted by prolonged pauses. *Nature Cell Biology*, *2*(3), 137–141. https://doi.org/10.1038/35004008

Wang, Lina, & Brown, A. (2010). A hereditary spastic paraplegia mutation in kinesin-1A/KIF5A disrupts neurofilament transport. *Molecular Neurodegeneration*, *5*(1), 52. https://doi.org/10.1186/1750-1326-5-52

Watila, M. M., & Balarabe, S. A. (2015). Molecular and clinical features of inherited neuropathies due to PMP22 duplication. In *Journal of the Neurological Sciences* (Vol. 355, Issues 1–2, pp. 18–24). Elsevier. https://doi.org/10.1016/j.jns.2015.05.037

Watson, J. D., & Crick, F. H. C. (1953). Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid. *Nature*, *171*(4356), 737–738. https://doi.org/10.1038/171737a0

Wilschanski, M., Yahav, Y., Yaacov, Y., Blau, H., Bentur, L., Rivlin, J., Aviram, M., Bdolah-Abram, T., Bebok, Z., Shushi, L., Kerem, B., & Kerem, E. (2003).

Gentamicin-Induced Correction of CFTR Function in Patients with Cystic Fibrosis and CFTR Stop Mutations . *New England Journal of Medicine*, *349*(15), 1433–1441. https://doi.org/10.1056/nejmoa022170

Woods, J. D., Khanlou, N., Lee, H., Signer, R., Shieh, P., Chen, J., Herzog, M., Palmer, C., Martinez-Agosto, J., & Nelson, S. F. (2020). Myopathy associated with homozygous PYROXD1 pathogenic variants detected by genome sequencing. *Neuropathology*, *40*(3), 302–307. https://doi.org/10.1111/neup.12641

Xu, Z., Cork, L. C., Griffin, J. W., & Cleveland, D. W. (1993). Increased expression of neurofilament subunit NF-L produces morphological alterations that resemble the pathology of human motor neuron disease. *Cell*, *73*(1), 23–33. https://doi.org/10.1016/0092-8674(93)90157-L

Yadav, P., Selvaraj, B. T., Bender, F. L. P., Behringer, M., Moradi, M., Sivadasan, R., Dombert, B., Blum, R., Asan, E., Sauer, M., Julien, J. P., & Sendtner, M. (2016). Neurofilament depletion improves microtubule dynamics via modulation of Stat3/stathmin signaling. *Acta Neuropathologica*, *132*(1), 93–110. https://doi.org/10.1007/s00401-016-1564-y

Ylikallio, E., Johari, M., Konovalova, S., Moilanen, J. S., Kiuru-Enari, S., Auranen, M., Pajunen, L., & Tyynismaa, H. (2014). Targeted next-generation sequencing reveals further genetic heterogeneity in axonal Charcot-Marie-Tooth neuropathy and a mutation in HSPB1. *European Journal of Human Genetics*, *22*(4), 522–527. https://doi.org/10.1038/ejhg.2013.190

Ylikallio, E., Pöyhönen, R., Zimon, M., De Vriendt, E., Hilander, T., Paetau, A., Jordanova, A., Lönnqvist, T., & Tyynismaa, H. (2013). Deficiency of the E3 ubiquitin ligase TRIM2 in early-onset axonal neuropathy. *Human Molecular Genetics*, *22*(15), 2975–2983. https://doi.org/10.1093/hmg/ddt149

Ylikallio, E., Woldegebriel, R., Tumiati, M., Isohanni, P., Ryan, M. M., Stark, Z., Walsh, M., Sawyer, S. L., Bell, K. M., Oshlack, A., Lockhart, P. J., Shcherbii, M., Estrada-Cuzcano, A., Atkinson, D., Hartley, T., Tetreault, M., Cuppen, I., Van Der Pol, W. L., Candayan, A., … Tyynismaa, H. (2017). MCM3AP in recessive Charcot-Marie-Tooth neuropathy and mild intellectual disability. *Brain*, *140*(8), 2093–2103. https://doi.org/10.1093/brain/awx138

Yuan, A., Sershen, H., Veeranna, Basavarajappa, B. S., Kumar, A., Hashim, A., Berg, M., Lee, J. H., Sato, Y., Rao, M. V., Mohan, P. S., Dyakin, V., Julien, J. P., Lee, V. M. Y., & Nixon, R. A. (2015). Neurofilament subunits are integral components of synapses and modulate neurotransmission and behavior in vivo. *Molecular Psychiatry*, *20*(8), 986–994. https://doi.org/10.1038/mp.2015.45

Yuan, Aidong, Rao, M. V., Kumar, A., Julien, J. P., & Nixon, R. A. (2003). Neurofilament Transport In Vivo Minimally Requires Hetero-Oligomer Formation. *Journal of Neuroscience*, *23*(28), 9452–9458. https://doi.org/10.1523/jneurosci.23-28-09452.2003

Yuan, Aidong, Rao, M. V., Sasaki, T., Chen, Y., Kumar, A., Veeranna, Liem, R. K. H., Eyer, J., Peterson, A. C., Julien, J. P., & Nixon, R. A. (2006). α-internexin is

structurally and functionally associated with the neurofilament triplet proteins in the mature CNS. *Journal of Neuroscience*, *26*(39), 10006–10019. https://doi.org/10.1523/JNEUROSCI.2580-06.2006

Yuan, Aidong, Rao, M. V., Veeranna, & Nixon, R. A. (2017). Neurofilaments and neurofilament proteins in health and disease. *Cold Spring Harbor Perspectives in Biology*, *9*(4). https://doi.org/10.1101/cshperspect.a018309

Yuan, Aidong, Sasaki, T., Kumar, A., Peterhoff, C. M., Rao, M. V., Liem, R. K., Julien, J. P., & Nixon, R. A. (2012). Peripherin is a subunit of peripheral nerve neurofilaments: Implications for differential vulnerability of cns and peripheral nervous system axons. *Journal of Neuroscience*, *32*(25), 8501–8508. https://doi.org/10.1523/JNEUROSCI.1081-12.2012

Yuan, Aidong, Sasaki, T., Rao, M. V., Kumar, A., Kanumuri, V., Dunlop, D. S., Liem, R. K., & Nixon, R. A. (2009). Neurofilaments form a highly stable stationary cytoskeleton after reaching a critical level in axons. *Journal of Neuroscience*, *29*(36), 11316–11329. https://doi.org/10.1523/JNEUROSCI.1942-09.2009

Yuan, Aidong, Veeranna, Sershen, H., Basavarajappa, B. S., Smiley, J. F., Hashim, A., Bleiwas, C., Berg, M., Guifoyle, D. N., Subbanna, S., Darji, S., Kumar, A., Rao, M. V., Wilson, D. A., Julien, J. P., Javitt, D. C., & Nixon, R. A. (2018). Neurofilament light interaction with GluN1 modulates neurotransmission and schizophrenia-associated behaviors. *Translational Psychiatry*, *8*(1). https://doi.org/10.1038/s41398-018-0194-7

Yum, S. W., Zhang, J., Mo, K., Li, J., & Scherer, S. S. (2009). A novel recessive NEFL mutation causes a severe, early-onset axonal neuropathy. *Annals of Neurology*, *66*(6), 759–770. https://doi.org/10.1002/ana.21728

Zhai, J., Lin, H., Julien, J. P., & Schlaepfer, W. W. (2007). Disruption of neurofilament network with aggregation of light neurofilament protein: A common pathway leading to motor neuron degeneration due to Charcot-Marie-Tooth disease-linked mutations in NFL and HSPB1. *Human Molecular Genetics*, *16*(24), 3103–3116. https://doi.org/10.1093/hmg/ddm272

Zhang, B., Tu, P. H., Abtahian, F., Trojanowski, J. Q., & Lee, V. M. Y. (1997). Neurofilaments and orthograde transport are reduced in ventral root axons of transgenic mice that express human SOD1 with a G93A mutation. *Journal of Cell Biology*, *139*(5), 1307–1315. https://doi.org/10.1083/jcb.139.5.1307

Zhu, Q., Couillard-Després, S., & Julien, J. P. (1997). Delayed Maturation of Regenerating Myelinated Axons in Mice Lacking Neurofilaments. *Experimental Neurology*, *148*(1), 299–316. https://doi.org/10.1006/exnr.1997.6654

9 CONTRIBUTIONS

9.1 Publication I

In this study I performed the genetic analysis of all patients, statistical analysis, review of background literature, prepared the figures, and wrote the original draft. Additionally, I participated in data and overall project management, and Sanger sequencing of patients and their families. I presented my work in national seminars.

9.2 Publication II

In this study I participated in the conceptualization, study design, review of literature, sequencing, and general project management. I wrote the original draft and prepared the figures.

9.3 Publication III

In this study I wrote the original draft and prepared the figures. I participated in data analysis, project conceptualization and experimental design. I did most of the laboratory work, including iPSC culture, iPSC-MN differentiation and optimization, immublotting, RT-qPCR, ICC and imaging, EM imaging and Sanger sequencing.

9.4 Publication IV

In this study I conceptualized the project and designed the experiments with my supervisors. I managed the project and performed majority of the experiments and data analysis with assistance from other authors. I wrote the original draft and prepared the figures. I optimized the iPSC-MN differentiation protocol, microfluidic culture of iPSC-MNs and organelle trafficking experiments. I did most of the laboratory work, including iPSC culture, iPSC-MN differentiation, immublotting, RT-qPCR, ICC and imaging, EM imaging, axotomy experiments, organelle tracking, NMDi treatment trials and Sanger sequencing.