Aus dem

Zentrum für Neurologie

Hertie Institut für klinische Hirnforschung

Abteilung Neurologie mit Schwerpunkt

Neurodegenerative Erkrankungen

Disease Gene Identification in Autosomal Recessive Hereditary Ataxias

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Medizin

der Medizinischen Fakultät
der Eberhard Karls Universität
zu Tübingen

vorgelegt von

Koch, Anika

2025

Dekan: Professor Dr. B. Pichler

1. Berichterstatter: Professorin Dr. R. Schüle-Freyer
2. Berichterstatter: Professor Dr. S. Ossowski

Tag der Disputation: 30.10.2025

Content

Table of Content III

List of Figures VI

List of Tables IX

List of Abbreviations XI

1 Introduction 1

1.1 Cerebellar ataxias 1
1.2 The role of the cerebellum 3
1.3 Diagnostics of cerebellar ataxias 7
1.4 Clinical presentation of hereditary cerebellar ataxias 8
1.5 Genetics of hereditary cerebellar ataxias 10
1.6 Primary research project and thesis objective 14

2 Materials and Methods 15

2.1 Materials 15
2.1.1 Devices, Software and Kits 15
2.1.2 Reagents for PCR, Agarose Gel Electrophoresis and Sequencing 16
2.1.3 PCR and sequencing primer 17
2.1.4 Cell culture media and supplements 18
2.2 Methods 19
2.2.1 In silico 19
2.2.1.1 Forming the cohort 19
2.2.1.2 NGS data analysis using the Genesis platform 21
2.2.1.3 Omim 29
2.2.1.4 ExAC and GnomAD 29
2.2.1.5 Constraint scores 29
2.2.1.6 Conservation scores 30
2.2.1.7 CADD Score 31
2.2.1.8 ACMG 31
2.2.1.9 Tissue expression 32
2.2.1.10 Interaction between known ataxia genes and novel candidate genes 32
2.2.1.11 Primer design 33
2.2.2 In vitro 33
2.2.2.1 Polymerase chain reaction 33

2.2.2.2 Touchdown PCR 34
2.2.2.3 Gel Electrophoresis 35
2.2.2.4 Sequencing PCR 36
2.2.2.5 Sanger Sequencing 36
2.2.2.6 Fibroblast culture and RNA Isolation 37
2.2.2.7 Cordycepin 37
2.2.2.8 Polyadenylation Assay 37
3 Results 40
3.1 Overview of cohort 40
3.2 Known genes 43
3.2.1 *AFG3L2* – Family 29 46
3.2.2 *PTRH2* – Family 53 49
3.2.3 *RNF216* – Family 58 51
3.2.4 *SBF1* – Family 59 54
3.2.5 *SEPSECS* – Family 81 56
3.2.6 *NGLY1* – Family 88 58
3.2.7 *MTPAP* – Family 16 60
3.3 Solved by other laboratories 63
3.3.1 *ATAD3A* – Family 26 64
3.3.2 *MT_ATP6* – Family 39 66
3.3.3 *FXN* – Family 69 68
3.4 New candidate genes 69
3.4.1 *GPAA1* – Family 43 and 62 73
4 Discussion 78
4.1 Known genes 83
4.1.1 *AFG3L2* – Family 29 83
4.1.2 *PTRH2* – Family 53 85
4.1.3 *RNF216* – Family 58 86
4.1.4 *SBF1* – Family 59 87
4.1.5 *SEPSECS* – Family 81 88
4.1.6 *NGLY1* – Family 88 91
4.1.7 *MTPAP* – Family 16 92
4.2 Solved by other laboratories 93
4.2.1 *ATAD3A* – Family 26 93
4.2.2 *MT_ATP6* – Family 39 93

4.2.3 *FXN* – Family 69 ... 94
4.3 New candidate genes ... 95
4.3.1 *GPAA1* – Families 43 and 62 ... 95
5 Abstract ... 99
6 Zusammenfassung ... 100
7 References ... 101
8 Erklärungen zum Eigenanteil der Dissertationsschrift ... 115
9 Acknowledgements ... 116

List of Figures

Figure 1: Categorisation of hereditary ataxias according to Palau and Espinós, 2006; Synofzik, Schöls and Riess, 2013; Taroni, Chiapparini and Mariotti, 2013; Cabal-Herrera et al., 2020. 1
Figure 2: Categorisation of acquired ataxias according to Nachbauer, Eigentler and Boesch, 2015. 2
Figure 3: Starting from the original database of 669 patients, systematic sorting and exclusion led to our final database of 125 patients. 19
Figure 4: The final dataset was analysed to determine the most likely inheritance pattern for each family. Ultimately, 103 families were selected to form the final cohort. 20
Figure 5: The resulting final cohort consisting of 103 families was subdivided into simplex and multiplex cases and with regard to consanguinity. 21
Figure 6: All 103 families forming the final cohort were first analysed for the presence of variants in already known ataxia-causing genes. 24
Figure 7: After exclusion of the families that were solved either by already known genes or by other laboratories, the remaining multiplex and simplex consanguine families formed the seed cohort. All simplex non-consanguine families and the PREPARE group formed the confirmation cohort. 28
Figure 8: Distribution of our 103 index cases into simplex and multiplex cases as well as consanguineous and non-consanguineous cases. 41
Figure 9: Procedure for searching in already known ataxia-causing genes. All variants found are checked for plausibility and then prioritised further. 43
Figure 10: This is an overview of the number of families that could be genetically solved by analysing known genes. 44
Figure 11: Pedigree of family 29. 46
Figure 12: Pedigree of family 53. 49
Figure 13: Pedigree of family 58. 51
Figure 14: Pedigree of family 59. 54
Figure 15: Pedigree of family 81. 56
Figure 16: Pedigree of family 88. 58
Figure 17: Pedigree of family 16. 60

Figure 18: Fragment analysis of the polyadenylation assay. There is no difference in the polyadenylation status between the affected patients and the healthy control patient. .. 62
Figure 19: After analysing all families for variants in already known ataxia causing genes, 3 more families were solved by other laboratories. 63
Figure 20: Pedigree of family 26. 64
Figure 21: Pedigree of family 39. 66
Figure 22: Pedigree of family 69. 68
Figure 23: Forming and analysing the seed cohort consisting of all multiplex and simplex consanguine cases using Genesis. 69
Figure 24: The search for new genes started with an analysis in the seed cohort in Genesis. The identified variants were then rechecked in the confirmation cohort. 70
Figure 25: Further in silico analysis and prioritisation led to the identification of 21 potential new candidate genes. 71
Figure 26: Pedigree of family 43. 73
Figure 27: Pedigree of family 62. 74
Figure 28: Sanger sequencing analysis of the affected homozygous index patient, one unaffected heterozygous family member and an unaffected wildtype control for each of the two families. 76
Figure 29: The surface abundance of FLAER, a marker for GPI anchor proteins, was measured on control (upper image) and patient (lower image) PBMCs by FACS analysis. Patient fibroblasts show a reduction of GPI anchor protein expression on the cell surface by ~30%. 77
Figure 30: AFG3L2 protein with variant distribution and their phenotypical association (red = optic atrophy, black = spinocerebellar ataxia 28, blue = spastic ataxia syndrome 5). Source: Reproduced with permission from Annals of Neurology, John Wiley and Sons, "ATPase Domain AFG3L2 Mutations Alter OPA1 Processing and Cause Optic Neuropathy", Taroni et al., 2020, https://doi.org/10.1002/ana.25723 84
Figure 31: Panel A is a sagittal cerebral magnetic resonance image showing diffuse cerebellar atrophy (arrow) and cortical atrophy. Panel B is a transverse image and shows multiple distinct and confluent hyperintensity foci in the white matter. Reproduced with permission from "Ataxia, Dementia, and Hypogonadotropism Caused by Disordered

Ubiquitination" by Margolin et al. 2013 (DOI: 10.1056/NEJMoa1215993), Copyright Massachusetts Medical Society. 86
Figure 32: Biosynthesis of Sec-tRNA. The upper arrow shows the synthesis in bacteria using the selenocysteine synthase. The lower arrow shows the synthesis in eukaryotes and archae using an intermediate step and the enzyme SEPSECS to synthesize Sec-tRNA. Reprinted from The American Journal of Human Genetics, Agamy et al., Mutations Disrupting Selenocysteine Formation Cause Progressive Cerebello-Cerebral Atrophy, 2010, https://doi.org/10.1016/j.ajhg.2010.09.007, with permission from Elsevier........ 88
Figure 33: Cerebral magnetic resonance images of a patient with cerebellar atrophy (b), compared to a control (a) and a PCH2A patient (c). Source: van Dijk, T., Vermeij, JD., van Koningsbruggen, S. et al. A SEPSECS mutation in a 23-year-old woman with microcephaly and progressive cerebellar ataxia. J Inherit Metab Dis 41, 897–898 (2018). https://doi.org/10.1007/s10545-018-0151-x, http://creativecommons.org/licenses/by/4.0/ 89
Figure 34: Sagittal (left) and coronar (right) cerebral magnetic resonance image of our index patient showing a significant atrophy of the cerebellum. There are no signs that would be indicative of ARSACS or leukoencephalopathy. 89
Figure 35: Tissue gene expression of GPAA1. It shows low tissue specificity (source: https://gtexportal.org/home/gene/GPAA1)... 95
Figure 36: Tissue expression of GPAA1. There is no specific tissue expression (source: https://www.proteinatlas.org/ENSG00000197858-GPAA1). 96
Figure 37: GPI anchored protein and its structural features. Reprinted from Journal of Lipid Research, Susanne Heider,John A. Dangerfield,Christoph Metzner et al., Biomedical applications of glycosylphosphatidylinositol-anchored proteins, 2016, https://doi.org/10.1194/jlr.R070201, http://creativecommons.org/licenses/by/4.0/ 96

List of Tables

Table 1: Devices used in this study. 15
Table 2: Software used in this study. 15
Table 3: Kits used in this study. 16
Table 4: Components and manufacturing companies for PCR, Agarose Gel Electrophoresis and Sequencing. 16
Table 5: Sequencing primer pairs, all primers were purchased from Integrated DNA Technologies (Coralville, Iowa, USA) 17
Table 6: Cell culture media and manufacturing companies. 18
Table 7: Supplements and manufacturing companies. 18
Table 8: Known ataxia genes (table provided by Prof. Dr. med. R. Schüle-Freyer via the European Reference Network for Rare Neurological Diseases (ERN RND), status as of 05/2017). 25
Table 9: ACMG classification point system to interpret variants (source: Tavtigian et al., 2020). 31
Table 10: PCR reaction mix. 33
Table 11: PCR protocol. 34
Table 12: tdPCR protocol. 34
Table 13: Components 2% agarose gel. 35
Table 14: Sequencing PCR reaction mix. 36
Table 15: Sequencing PCR protocol. 36
Table 16: Poly(G/I)-tail-mix used in the polyadenylation assay. 38
Table 17: Reverse transcription mix used in the polyadenylation assay. 38
Table 18: PCR mix used in the polyadenylation assay. 39
Table 19: PCR protocol used in the polyadenylation assay. 39
Table 20: Overview of the segregating variants in known ataxia genes. 45
Table 21: Overview of the AFG3L2 variants. 46
Table 22: Overview of the PTRH2 variant. 49
Table 23: Overview of the RNF216 variant. 51
Table 24: Overview of the SBF1 variant. 54
Table 25: Overview of the SEPSECS variant. 56
Table 26: Overview of the NGLY1 variant. 58

Table 27: Overview of the MTPAP variant.. 60
Table 28: Overview of the variants in known ataxia genes that were found by other laboratories. 63
Table 29: Overview of the ATAD3A variant. 64
Table 30: Overview of the MT_ATP6 variant. 66
Table 31: Overview of the 21 new potential ataxia-causing genes identified in silico in the seed cohort. 71
Table 32: Overview of the GPAA1 variants in family 43 and 62. 75
Table 33: Genetic analysis in family 43 and family 62.. 76

List of Abbreviations

Abbreviation	Description
ACMG	American College of Medical Genetics and Genomics
AFG3L2	ATPase Family Gene 3 Like 2
ARSACS	Autosomal Recessive Spastic Ataxia of Charlevoix-Saguenay
ATAD3A	ATPase Family, AAA Domain-Containing Member 3A
cDNA	Copy DNA
ClinVar	Clinically Relevant Variant
cMRI	Cerebral Magnetic Resonance Imaging
ddH_2O	Double-Distilled Water
CADD	Combined Annotation-Dependent Depletion
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic Acid
dNTP	Deoxynucleotide Triphosphate
DP	Depth of Data
EDTA	Ethylenediaminetetraacetic Acid
EtOH	Ethanol
EVS	Exome Variant Server
ExAC	Exome Aggregation Consortium
FACS	Fluorescence-Activated Cell Sorting
FBS	Fetal Bovine Serum
FLAER	Fluorescein-Labeled Proaerolysin
FW	Forward
FXN	Frataxin
GnomAD	Genome Aggregation Database
GPAA1	Glycosylphosphatidylinositol Anchor Attachment Protein 1
GPI	Glycosylphosphatidylinositol
GQ	Genotype Quality
GTEx	Genotype-Tissue-Expression

HIPPIE	High-throughput Identification Pipeline for Promoter Interacting Enhancer Elements
INDEL	Insertion or Deletion
MAF	Minor Allele Frequency
MgCl2	Magnesium Chloride
mRNA	Messenger Ribonucleic Acid
MT_ATP6	Mitochondrial ATP Synthase 6
MT-PAP	Mitochondrial Poly(A) Polymerase
NaAc	Sodium Acetate
NCBI	National Center for Biotechnology Information
NGLY1	N-Glycanase 1
NGS	Next Generation Sequencing
NIH	National Institutes of Health
Omim	Online Mendelian Inheritance in Man
OPA1	OPA1 Mitochondrial Dynamin-Like GTPase
PBMC	Peripheral Blodd Monocytes
PBS	Dulbecco's Phosphate Buffered Saline
PCH2A	Pontocerebellar Hypoplasia 2A
PCR	Polymerase Chain Reaction
PHAST	Phylogenetic Analysis with Space/Time Models
PhastCons	PHAST Conservation Score
PhyloP	Phylogenetic P-values
pLI	Probability of Loss of Function Intolerance
PTRH2	Peptidyl-tRNA-Hydrolase 2
QUAL	Quality
RNA	Ribonucleic Acid
RNF216	Ring Finger Protein 216
RT	Reverse Transcriptase
RV	Reverse
SARA	Scale for Assessment and Rating of Ataxia
SBF1	SET Binding Factor 1
SEPSECS	O-Phosphoserine-tRNA-Selenocysteine-tRNA-Synthase

SNV	Single Nucleotide Variant
UCSC	University of California, Santa Cruz
WES	Whole Exome Sequencing
WGS	Whole Genome Sequencing

1 Introduction

1.1 Cerebellar ataxias

Ataxia is a derived term from the ancient Greek word ἀταξία which translates to disorder or absence of order. (Beal, Lang and Ludolph, 2005)

Cerebellar ataxias are a clinically and genetically very heterogenous group of neurodegenerative disorders. In general, cerebellar ataxias can be divided into hereditary, acquired and sporadic ataxias.

Hereditary ataxias are caused by inherited genetic variant. As seen in Figure 1 hereditary ataxias can be subdivided according to the inheritance into autosomal dominant, autosomal recessive, X-linked and mitochondrial. (Palau and Espinós, 2006; Synofzik, Schöls and Riess, 2013; Taroni, Chiapparini and Mariotti, 2013; Cabal-Herrera *et al.*, 2020)

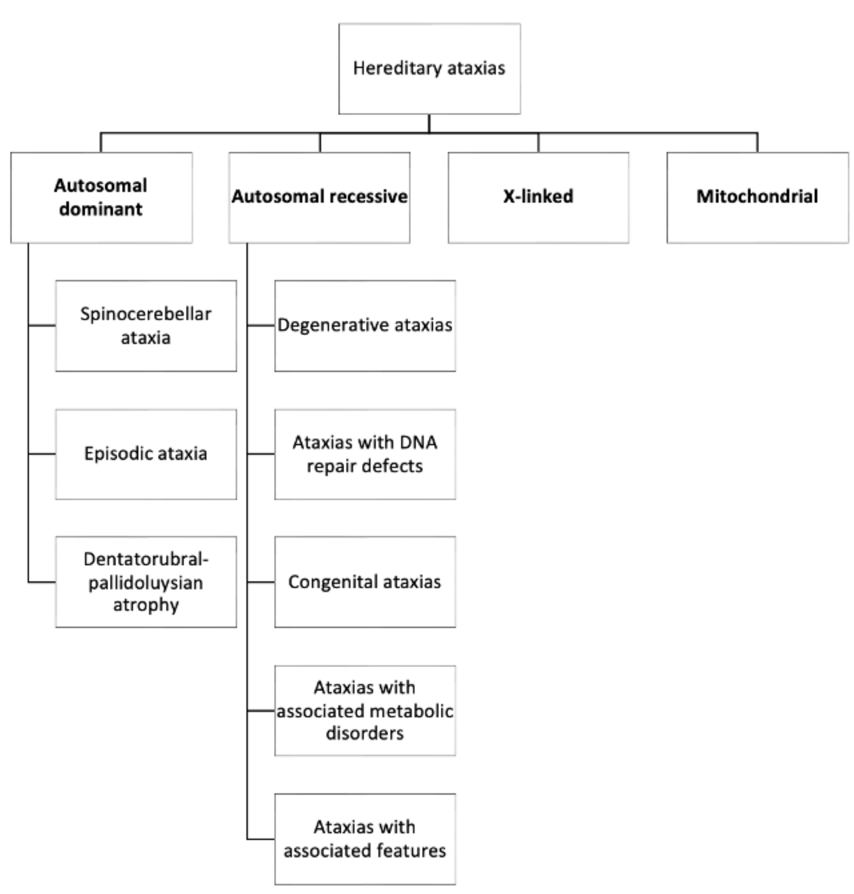

Figure 1: Categorisation of hereditary ataxias according to Palau and Espinós, 2006; Synofzik, Schöls and Riess, 2013; Taroni, Chiapparini and Mariotti, 2013; Cabal-Herrera et al., 2020.

The average prevalence of autosomal dominant cerebellar ataxias is 2.7/100.000 with Spinocerebellar ataxia type 3 (SCA3), also known as Machado-Joseph disease, being the most common form. In autosomal recessive cerebellar ataxias, the average prevalence is 3,3:100.000. Friedreich ataxia is reported to be the most common form. (Ruano *et al.*, 2014; Pilotto and Saxena, 2018)

Regarding X-linked cerebellar ataxia the most common form is the Fragile-X-associated tremor/ataxia-Syndrome with a prevalence of 1:4.000 in men and 1:8.000 in women. (Hagerman, 2008)

With mitochondrial ataxias the prevalence is estimated to be 1:5000. (Gorman *et al.*, 2015)

As shown in Figure 2 acquired ataxias are diverse with a broad spectrum of possible causes. (Nachbauer, Eigentler and Boesch, 2015) They can occur at any age, but the probability of developing an acquired ataxia increases with age. It is particularly important to conduct a careful search for acquired ataxia, as there are more treatment options available for this condition than for genetic and sporadic degenerative ataxias.

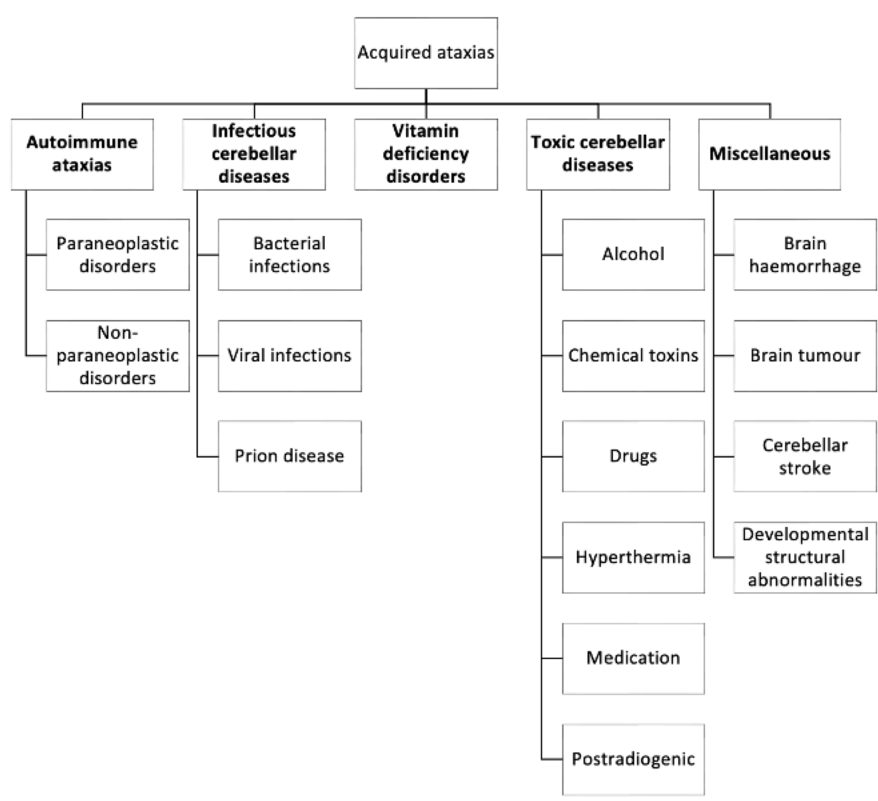

Figure 2: Categorisation of acquired ataxias according to Nachbauer, Eigentler and Boesch, 2015.

Sporadic degenerative ataxias and their diagnosis pose a major challenge. All non-genetic and genetic causes of ataxia must be considered to rule out hereditary and acquired ataxias.

Sporadic degenerative ataxias include the cerebellar type of multiple system atrophy and sporadic adult-onset ataxia of unknown aetiology. (Klockgether, 2010; Gaalen and Warrenburg, 2012; Lin and Kuo, 2023)

1.2 The role of the cerebellum

Cerebellar ataxias encompass a broad spectrum of neurodegenerative diseases whose common feature is the progressive dysfunction in the cerebellum. The characteristic symptoms of affected individuals range from gait instability, dysarthria and oculomotor disorder to cognitive impairments and systemic manifestations caused by degeneration or abnormal development of the cerebellum and its connections.

The cerebellum is responsible for regulating motor function, coordinating the body's complex movements through somatosensory feedback and motor learning processes as well as nonmotor functions such as abstract thinking, fluency of speech and social cognition. In cases of cerebellar damage, movements, posture, stance and gait tend to become unsteady, and it becomes difficult to perform precise movements. Speech disorders and hypotonia of the muscles also usually occur as well as a disturbed mentalising process and an impaired social cognition and affection. (Buckner, 2013; Ahmadian *et al.*, 2019; Van Overwalle *et al.*, 2020; Rudolph *et al.*, 2023)

Anatomically the cerebellum is divided into two hemispheres and a vermis. Microscopically, it can be further subdivided into the cerebellar cortex and the cerebellar medulla with the cerebellar nuclei.

The cerebellar cortex is again divided vertically into the following three zones: vestibulocerebellum, spinocerebellum and pontocerebellum. (Winser *et al.*, 2018)

Each zone mediates specific functions:

Vestibulocerebellum:

The vestibulocerebellum is phylogenetically the oldest part of the cerebellum and

regulates balance as well as postural and oculomotor functions. It receives primary vestibular afferents via the vestibular nuclei, which convey information about position and acceleration. It also receives visual input from the pretectum and the visual cortex. Furthermore, it receives somatosensory information regarding the body's position.

Processing these three types of input the vestibulocerebellum maintains the balance of the body by tensing or relaxing the trunk and leg muscles in a compensatory manner when movement of a limb causes impairment.

The medial part of the vestibulocerebellum coordinates balance by controlling the trunk muscles and the extremity extensors via the nucleus vestibularis lateralis as well as the vestibulospinal tracts. Therefore, lesions in this part lead to balance disorders. Patients have difficulties standing upright (stance ataxia) or sitting (truncal ataxia).

The lateral parts of the vestibulocerebellum regulate head and eye movements via the nucleus vestibularis medialis. Lesions in this part lead to problems fixating the eyes on a point. For instance, if an object moves continuously in the direction of the damaged side, it can only be followed by sudden eye movements (saccades). The disturbance of the fixation of the gaze on a stable point leads to spontaneous nystagmus, whereby the eyes wander uncontrollably before being brought back by saccades, and disturbance of the vestibulo-ocular reflex, which normally leads to gaze stabilization by turning the eyes in proportion to the head movement in the opposite direction. (Park *et al.*, 2021)

Spinocerebellum

The spinocerebellum controls muscle tone. It consists of two parts: The medial vermis and the paramedian zones. This area of the cerebellum controls the precise coordination of muscle activity during movement. Therefore, the spinocerebellum is responsible for continuously comparing the desired position with the actual position. Furthermore, it can also compensate for balance and posture disorders caused by movement.

The spinocerebellum receives afferents directly from the spinal cord. Lesions in the spinocerebellum result in disorders affecting motor functions such as target and support. Movements can no longer be continuously corrected by processing sensory feedback from muscles and joints. In the case of dysmetria, patients either overshoot or fail to reach the target. Problems can also occur when performing rapid alternating movements, such

as screwing in a light bulb, and are referred to as dysdiadochokinesia. Trembling during a targeted movement is referred to as intention tremor. This form of tremor typically increases as the target is approached, resulting from excessive and slow corrective movements. Rather than initiating corrections based on the anticipated movement trajectory of the efferent copy before execution, the cerebellum can only make subsequent corrections based on incoming afferents.

Pontocerebellum

In evolutionary terms, the pontocerebellum is the youngest part of the cerebellum and corresponds to the two cerebellar hemispheres. It communicates almost exclusively with the cerebral cortex and receives no sensory afferents. It receives inputs from the prefrontal, premotor and supplementary motor cortex areas via the pons. Additionally, it receives motor efferent copies from the primary motor cortex and the spinocerebellum via the inferior olive.

The main functions of the pontocerebellum are motor learning and creating complex movement sequences, such as speech motor skills or complicated hand movements like playing music. These sequences require rapid and precise movements from many muscles, making feedback regulation difficult. The pontocerebellum acts as a procedural memory, enabling implicit learning. After extensive training, it stores and enables very complicated automated movement sequences so that they no longer require conscious thought, with the individual steps largely performed without sensory feedback or conscious control. This distinguishes the pontocerebellum functionally from the other two parts of the cerebellum, which use sensory feedback to mediate movement sequences and corrections.

Failure symptoms in the pontocerebellum lead to impairments in complex, learnt movements, resulting in asynergy, dysmetria, dysdiadochokinesia, action tremor and a loss of coordination in the individual muscle movements of a sequence. Unilateral lesions lead to limb ataxia. If speech motor skills are affected, the result is dysarthria with unclear, slurred or choppy speech. Asynergy can also lead to the decomposition of movement sequences, whereby the components of a movement are performed one after the other rather than simultaneously. (Dietrichs, 2008)

Horizontally the cerebellar cortex consists of three cell layers and comprises five types of nerve cells (stellate cells, basket cells, golgi cells, granule cells and Purkinje cells), some of which have variable synapses that enable motor learning and synaptic plasticity.

The molecular layer contains dendrites of Purkinje cells, axons of the granule cells (parallel fibres), climbing fibres and interneurons, especially stellate cells and basket cells.

The Purkinje cell layer consists of the large, inhibitory Purkinje cells, which are the only output structure of the cerebellar cortex, and Bergmann glia.

The granular cell layer contains small, excitatory granule cells and Golgi cells (inhibitory interneurons).

These three layers are functionally closely intertwined in a stereotypical manner to perform the cerebellum's corrective and learning functions at a cellular and molecular level.

The calculations and output signals in the cerebellum are essentially based on the targeted and modifiable inhibition of existing activity, not on the generation of new signals. The two main excitatory inputs to the cerebellar cortex are provided by the mossy and climbing fibres, which also stimulate the cerebellar nuclei via axon collaterals. The local interconnection of the five cell types enables context-dependent processing of the original input signal within the cerebellar cortex through mechanisms such as divergence, spatial and temporal contrast enhancement and synaptic plasticity.

The GABAergic Purkinje cells form the central integration station of the cerebellar cortex. They are the only projection neurons in this region and selectively inhibit the stimulation of certain populations of neurons in the cerebellar nuclei. The remaining excitation pattern of the cerebellar nuclei’s projection neurons forms the actual integrative output signal of the cerebellum.

In hereditary cerebellar ataxias the three layers are affected to varying degrees, resulting in impaired coordination of movement and balance coordination. (Liss and Kätzel, 2019)

In the molecular layer dysfunction of interneurons can interfere with the fine modulation of Purkinje cell activity. In spinocerebellar ataxia type 1, it has been demonstrated that

the hyperactivity of the interneurons in the molecular layer results in the degeneration of Purkinje cells, which impairs synaptic integration and inhibitory control. (Pilotto *et al.*, 2023)

Purkinje cells in the Purkinje cell layer are particularly susceptible to genetic ataxia. They project GABA to the cerebellar nuclei, thereby controlling motor fine-tuning. Their degeneration is crucial for the development of symptoms. Failure of these cells leads to uncoordinated movements. (Becker *et al.*, 2009) Variants in the ATXN1, ATXN2 or ITPR1 genes have been shown to directly impair the function or survival of Purkinje cells. (Kashiwabuchi *et al.*, 1995; Serra *et al.*, 2006; Kasumu and Bezprozvanny, 2012; Paul and Limaiem, 2025)

In the granular layer, granule cells receive input from the mossy fibres from the spinal cord and brainstem, transmitting them to the Purkinje cells via parallel fibres. Disorders in this layer (e.g. due to disturbed synapse formation or cell loss) can impair information processing throughout the cerebellar circuit. In some hereditary ataxias, such as Friedreich's ataxia, degeneration of the granule cells or their afferent pathways can also occur. (Pascual-Castroviejo *et al.*, 1994; Simon *et al.*, 2004)

1.3 Diagnostics of cerebellar ataxias

When it comes to diagnosing cerebellar ataxias, taking a detailed medical history and conducting a thorough clinical examination are the most important steps. It is particularly important to take the family history as accurately and comprehensively as possible, given that comparable movement disorders in relatives might be a key indicator of genetic ataxia. However, a negative family history does not rule out the possibility of genetic ataxia. The clinical neurological examination includes testing for the presence of ataxia-typical symptoms of gait and stance ataxia, tremor, dysarthria and nystagmus. Coordination and reflex exercises are also carried out.

Even though in most cases it is already possible to differentiate clinically between ataxia and a focal cerebellar disease (tumor, abscess, multiple sclerosis and others), a clear distinction can only be made with an MRI of the brain, which is usually the next diagnostic step. In some cases of ataxia, a diagnosis can be made based on the clinical

presentation and MRI scan results (e.g. leptomeningeal hemosiderosis). If a diagnosis is not possible at this stage, biochemical and molecular genetic tests are performed. Blood and cerebrospinal fluid can be examined for evidence of acquired ataxias. If biochemical screening does not lead to a definitive diagnosis, genetic testing will follow. This is initially based on family history and the most likely mode of inheritance. If a definitive diagnosis cannot be made at this stage either, gene panel examination, whole exome sequencing or whole genome sequencing is carried out. (Jacobi and Minnerop, 2021)

1.4 Clinical presentation of hereditary cerebellar ataxias

Hereditary autosomal recessive cerebellar ataxias are often multisystemic diseases that can involve the central as well as the peripheral nervous system. The huge variety of partially overlapping symptoms and large genetic heterogeneity complicate finding the correct diagnosis. (Manto and Marmolino, 2009)

Cerebellar ataxia is a neurological syndrome caused by dysfunction in the cerebellum or its afferent and efferent connections. The condition manifests as a variety of motor and, occasionally, cognitive symptoms, which can be traced back to the cerebellum's central role in coordinating, fine-tuning, and planning movement sequences.

A central feature is broad-based gait ataxia, characterised by an unsteady, staggering walk. The stride length is irregular, with frequent stumbles and falls occurring. An unsteady stance with a tendency to sway or fall without external influence may also occur, particularly when the eyes are closed. Target movement disorders may also be present. Dysmetria is evident in voluntary movements and is often accompanied by hypermetria, hypometria and an intention tremor.

Dysdiadochokinesia is observed when the coordination of rapid, alternating movements is impaired. Cerebellar dysarthria can occur and is also known as ataxic dysarthria. Speech is slow and choppy, with irregular accentuation of syllables and explosive vocalisations. (Manto and Mariën, 2015)

The cerebellum is also involved in coordinating oculomotor functions. Consequences of disorders in this area include nystagmus, saccades and disorders of eye tracking, as well as impaired fixation and gaze stabilisation.

Decreased muscle tone at baseline, manifested as muscular hypotonia, may be observed. This results in the extremities appearing flaccid and increased pendular movement during reflexes.

Non-motor symptoms that may be observed include affective symptoms such as apathy, depression, and disinhibition, as well as executive dysfunction involving reduced cognitive flexibility, slower thinking, and a decreased problem solving and self regulation abilities. (Schmahmann and Sherman, 1998; Schmahmann, 2004)

Attention and working memory disorders may also be present. Visual-spatial impairments manifest as a disturbed sense of spatial orientation. Linguistic changes do not manifest as classic aphasia, but rather as reduced verbal fluency, limited sentence structure, and an imprecise choice of words. Deficits in understanding social contexts, empathy, and recognising emotions may also be present. (Hoche *et al.*, 2016)

To assess disease severity and progression, the International Cooperative Ataxia Rating Scale (ICARS) was established. It consists of the four parts postural and stance disturbances, limb movements disturbances, speech disorders and oculomotor disorders. The highest possible score is 100. (Trouillas *et al.*, 1997)

Later, the scale for assessment and rating of ataxia (SARA) was introduced. It is easier to administer, and its completion requires less than 15 minutes. The scale is divided into the following eight categories: gait, stance, siting, speech disturbance, finger chase, nose-finger-test, fast alternating hand movements and heel-shin slide. A total score between 0 to 40 is possible, with 0 being no ataxia and 40 being the most severe form of ataxia. (Schmitz-Hübsch *et al.*, 2006)

1.5 Genetics of hereditary cerebellar ataxias

The most common features in hereditary cerebellar ataxias are an impaired speech, limb coordination, poor balance and a gait disorder. Frequently observed neurological signs and symptoms include cerebellar oculomotor disturbances with gaze evoked nystagmus, saccadic pursuit and dysmetric saccades, and limb ataxia with dysmetria and dysdiadochokinesis. (Hersheson, Haworth and Houlden, 2012; Vedolin *et al.*, 2013; Arslan *et al.*, 2017)

The molecular diagnosis of hereditary cerebellar ataxias relies mostly on genetic testing. With more than 150 known causative genes, the genetic etiology of ataxias is extremely heterogeneous. Traditionally, hereditary ataxias are classified according to their mode of inheritance. The inheritance patterns include autosomal dominant and recessive, mitochondrial and X-linked inheritance.

Observed variant mechanisms are also heterogeneous and variants leading to a gain or loss of function have been described. (Hersheson, Haworth and Houlden, 2012)

Genetic mechanisms can be divided into four main categories.

1. Trinucleotide repeat expansions:
   This usually affects CAG, GAA or CTG repeats. Many autosomal dominant ataxias are caused by CAG repeat expansions. The main pathophysiological mechanism of these expansions is the toxic effect of abnormally elongated polyamino acid chains. In the case of CAG repeat expansions, polyglutamine-mediated toxicity occurs. Machado-Joseph disease (spinocerebellar ataxia 3) is an example of this. (Durr, 2010; Matilla-Dueñas *et al.*, 2014)
   Intronic repeat expansions also exist, as in the case of Friedrich ataxia. In this condition, GAA expansions occur in intron 1. This ultimately leads to reduced frataxin expression, resulting in mitochondrial iron accumulation and oxidative damage to nerve cell populations. (Koeppen, 2011)
2. Loss-of-function variants:
   Autosomal recessive ataxias are typically caused by variants that result in the complete or partial loss of function of the encoded protein. This can affect functions associated with mitochondria, metabolism or DNA repair.

In the case of cerebellar ataxia with CoQ10 deficiency, for example, the mitochondrial respiratory chain is disrupted, resulting in a deficiency of ATP and subsequent neuronal cell death. (Lopriore *et al.*, 2024)

A metabolic disorder is present in ataxia with vitamin E deficiency. In this case, a variant in the *TTPA* gene leads to impaired vitamin E transport, resulting in oxidative stress and nerve cell damage. (Thapa *et al.*, 2022)

Ataxia-telangiectasia is characterised by a variant in the *ATM* gene that leads to a defect in DNA double-strand break repair. (Amirifar *et al.*, 2019)

3. Gain-of-function variants:

   In some cases, a variant can result in the production of proteins that impair cell function in a toxic way. One example of this is dentatorubral-pallidoluysary atrophy. A variant in the *ATN1* gene causes the production of a mutated atrophin-1 protein that has an abnormally long polyglutamine tract. This protein then accumulates in the cells, disrupting normal neuronal processes and ultimately resulting in neuronal dysfunction and cell death. (Sato *et al.*, 1999)

4. Mitochondrial variants:

   These are mtDNA variants, which are usually inherited from the mother and include point variants or deletions of mitochondrial DNA. One example is MELAS (Mitochondrial Encephalomyopathy with Lactic Acidosis and Stroke-like episodes) syndrome, which is caused by a variant in the *MT-TL1* gene and leads to mitochondrial encephalopathy with ataxia. There are also nuclear genes that have a mitochondrial function, such as *POLG*. These genes code for proteins involved in mitochondrial replication or repair. Variants in these genes can result in ataxia, epilepsy, myopathy, and sensory neuropathy. (El-Hattab *et al.*, 2015; Vernon and Bindoff, 2018)

Typically, variants in ataxia-causing genes result in a progressive atrophy of the cerebellum and Purkinje Cell degeneration leading to the atactic phenotype. (Sakaguchi *et al.*, 1996; Lim *et al.*, 2006; Huang and Verbeek, 2019)

Despite the large number of known ataxia genes, nearly 30% of patients with hereditary ataxia remain without a genetic diagnosis. (Smeets and Verbeek, 2014)

On the one hand, unusual types of variants, e.g. structural variants, repeat expansions or variants located in non-coding regions that are challenging to detect even using whole genome sequencing contribute to this diagnostic gap. On the other hand, there are still unknown genes that cause ataxias.

Molecular genetic diagnostics play a central role in clarifying hereditary ataxias. Due to enormous genetic heterogeneity and overlapping clinical phenotypes, targeted genetic analysis is essential for precise diagnosis and prognosis assessment.

In order to genetically diagnose a person with atactic symptoms there are different possible initial situations. If a genetic variant is already known in this family, the patient can be tested for the presence of this very variant for example via Polymerase Chain Reaction and subsequent Sanger sequencing. The situation is quite different if the patient is the first to have the disease in the family or if there are several family members with the disease but no known genetic variant.

A step-by-step procedure has been established for genetic diagnostics. After clinical assessment and taking a family history, a targeted repeat-expansion analysis can be performed quickly and cost-effectively, but this does not cover other variant types or rare cerebellar ataxias. The minimum standard for molecular genetic diagnostics is the performance of a comprehensive ataxia-specific NGS panel. If this does not lead to a genetic diagnosis, then whole-exome or whole-genome sequencing can be performed. Whole-exome sequencing is an effective procedure to sequence all protein-coding regions of genes in a genome. The human exome constitutes about 1% of the human genome with approximately 180.000 exons. Therefore, it is considerably faster and cheaper than whole-genome sequencing. But in case of variants that are not located in protein-coding regions whole-genome sequencing is indispensable. In both procedures a biochemical analysis needs to be performed with this huge quantity of data. Candidate variants can be further analysed via Polymerase Chain Reaction (PCR) and Sanger sequencing of additional healthy or affected family members. (Ng *et al.*, 2009, 2010)

Genotype-phenotype correlations are a central topic of neurogenetic research, as they can be used to predict disease progression, age of onset and additional symptoms based on genetic findings.

Polyglutamine ataxias are a well-characterised group with relatively clear correlations between repeat length, disease onset and clinical symptoms. In spinocerebellar ataxia type 1, for example, the typical phenotype consists of ataxia, dysarthria and pyramidal tract signs, and it is known that the longer the repeat in the ATXN1 gene, the earlier the disease will manifest and the more severe it will be. (Zühlke *et al.*, 2002)

In the case of Friedreich's ataxia, the length of the GAA repeats in intron 1 of the FXN gene also determines the phenotype. If fewer than 300 repeats are present, the disease will have a later onset and a milder course. Conversely, if more than 700 GAA repeats are present, an earlier onset of the disease, rapid progression, and often cardiomyopathy can be expected. (Metz *et al.*, 2013; Lecocq *et al.*, 2016)

In ataxia-telangiectasia, the type of variant significantly influences both the age at which symptoms first appear and the extent to which multiple systems are affected. The classic form of ataxia-telangiectasia with early onset, pronounced immunodeficiency and an increased risk of malignancy is seen if there is a null variant (frameshift or nonsense) in the ATM gene, which leads to complete loss of function of the gene. This form is characterised by ataxia, oculocutaneous telangiectasia, and an increased susceptibility to infections. However, if a missense variant is present, a degree of residual enzyme activity may result in a mild form of the disease presenting with isolated cerebellar ataxia and occasionally parkinsonism, but with late onset. (Taylor and Byrd, 2005)

Knowledge of the exact genetic variant is exceptionally important in many regards. In most cases it takes a lot of years sometimes even decades to detect the underlying cause of a disease. This is very exhausting for both the patient and their relatives. Finding the reason that explains the patients' symptoms is a milestone towards acceptance and understanding of the disease. It furthermore is the foundation for more diagnostics towards previously undetected symptoms and for preventing complications. The genetic diagnosis is essential for physical medicine, rehabilitation and a possible cure. Moreover, it has an intrafamilial importance regarding the health of other family members and their family planning.

1.6 Primary research project and thesis objective

The central goal of this research project is the identification of novel ataxia gene candidates to increase the diagnostic rate in a cohort of previously undiagnosed hereditary ataxia cases compatible with autosomal recessive inheritance of ataxia. Towards this goal, we have defined the following specific objectives:

1. Compilation of the cohort: To compile the demographic and clinical information of unsolved cases, complete missing pedigrees, check sample availability and assemble NGS datasets for all unsolved families.
2. Re-analysis of NGS datasets for variants in known genes: To identify candidate variants in known ataxia genes using appropriate predefined filter settings for known ataxia genes, perform segregation studies, establish a genetic diagnosis and re-assess the family phenotype.
3. Identification of novel ataxia gene candidates: To identify candidate variants in potential novel disease genes based on appropriate filter settings and provide rich variant annotation allowing to prioritize candidate variants based on segregation in several families, predicted delirious effect on protein function, functional pathway annotation and expression profiles.

2 Materials and Methods

2.1 Materials

2.1.1 Devices, Software and Kits

Ethics vote reference number: 598/2011BO1

Table 1: Devices used in this study.

Device	Manufacturer
Centrifuge 5810R/5425	Eppendorf AG, Hamburg, Germany
Gel Documentation Chamber V02 9119	Viber Lirmat, Eberhardzell, Germany
Genetic Analyzer 3130xl	Applied Biosystems, Foster City, USA
Nanodrop ND1000	Peqlab Biotechnologie, Erlangen, Germany
T100^{TM} Thermal Cycler	Bio-Rad Laboratories, Hercules, USA

Table 2: Software used in this study.

Software	Company
3130xl Series DataCollection 4	Applied Biosystems, Foster City, USA
Bio-Capt 12.6 Software	Viber Lirmat, Eberhardzell, Germany
Nanodrop 1000 Software	Peqlab Biotechnologie, Erlangen, Germany
SeqA6 Sample manager	Open source
Staden Package (Pregap4)	Open source
Genesis	Open source
UCSC Genome Browser	Open source
Primer3web	Open source

Table 3: Kits used in this study.

Kits	Manufacturer
BigDye™ Terminator v3.1 Cycle Sequencing Kit	Life Technologies GmbH, Darmstadt, Germany
Poly(A) Tail Length Assay Kit; 76455	Affymetrix, Inc., Cleveland, USA
QIAquick Gel Extraction Kit; 28706	Qiagen, Hilden, Germany

2.1.2 Reagents for PCR, Agarose Gel Electrophoresis and Sequencing

Table 4: Components and manufacturing companies for PCR, Agarose Gel Electrophoresis and Sequencing.

Reagent	Company
Colorless GoTaq Reaction Buffer 5x	Promega, Madison, WI, USA
dNTPs (10 mM)	Life Technologies, Thermo Fisher Scientific, Waltham, USA
Formamide	Life Technologies, Thermo Fisher Scientific, Waltham, USA
Gel loading buffer	Life Technologies, Thermo Fisher Scientific, Waltham, USA
Gene Ruler DNA Ladder Mix	Life Technologies, Thermo Fisher Scientific, Waltham, USA
G2 GoTaq DNA Polymerase	Promega, Madison, WI, USA
Green GoTaq Reaction Buffer 5x	Promega, Madison, WI, USA
Midori Green Advance	Biozym, Vienna, Austria
SeaKem LE Agarose	Lonza, Basel, Switzerland
Tris-borate EDTA buffer pH=8 0,5x (TBE buffer)	Lonza, Basel, Switzerland
ddH_2O	Merck

2.1.3 PCR and sequencing primer

Table 5: Sequencing primer pairs, all primers were purchased from Integrated DNA Technologies (Coralville, Iowa, USA)

Name	Position	Primer sequences (5‘-3‘)
ATPase Family Gene 3 Like 2 (*AFG3L2*)	chr18:12340273-12340510	Fw: TCTTCAGAGACTCGACCACC Rv: TGCGGCTGTTGTGCTTATTT
ATPase Family Gene 3 Like 2 (*AFG3L2*)	chr18:12351147-12351394	Fw: CCAGTTTTAGCGGTCGGAGA Rv: ATGTCGTCATTTTGGCCGG
Glycosylphosphatidylinositol Anchor Attachment Protein 1 (*GPAA1*)	chr8:145139364-145139565	Fw: CAGACACTGCTGCTCATGGT Rv: CTCCAGCTAGACCCCATCCT
Glycosylphosphatidylinositol Anchor Attachment Protein 1 (*GPAA1*)	chr8:145138706-145138919	Fw: TGACTCTACCAACAGCCAGG Rv: CTTCAAGCCAAGCCTCAGTG
Mitochondrial Poly(A) Polymerase (*MTPAP*)	chr10:30629201-30629381	Fw: GGCAACTGATTACTTGACCGT Rv: AGGGTCTCTATGCTGTCGTA
N-Glycanase 1 (*NGLY1*)	chr3:25770619-25770770	Fw: GTTACCATGTGCCAGTCTGT Rv: ATTTCTAAACAGCTCCACCTTTG
Peptidyl-tRNA-Hydrolase 2 (*PTRH2*)	chr17:57774979-57775208	Fw: GAGCTTTGACCACCACCTTG Rv: GACGAGCAAGACACACACAG
Peptidyl-tRNA-Hydrolase 2 (*PTRH2*)	chr17:57775226-57775377	Fw: CCCAAAGCATACTCGAAGGC Rv: ACTGCCGGATAGTGATAACTGT
Ring Finger Protein 216 (*RNF216*)	chr7:5752301-5752460	Fw: AATGGGGAAGAGTTGGCAGA Rv: GTTCGTTCCCAACCAGTGAG

Ring Finger Protein 216 (*RNF216*)	chr7:5780781-5780931	Fw: AGAAGAGGGGCCTGAAATCC Rv: CCAACAGCCCCGTGAAATAA
O-Phosphoserine-tRNA-Selenocysteine-tRNA-Synthase (*SEPSECS*)	chr4:25125661-25125833	Fw: ACTTCGTTCTTTTCTTACTGCCT Rv: GGTCCATGCAAACTGTGAGT
O-Phosphoserine-tRNA-Selenocysteine-tRNA-Synthase (*SEPSECS*)	chr4:25125491-25125837	Fw: TGCTTGCTTGTACTACAGCCT Rv: CTTGGGTCCATGCAAACTGTG

2.1.4 Cell culture media and supplements

Table 6: Cell culture media and manufacturing companies.

Medium	Company
Dulbecco's Modified Eagle's Medium (DMEM)	Life Technologies GmbH, Darmstadt, Germany
Dulbecco's Phosphate Buffered Saline (PBS)	PAA laboratories GmbH, Pasching, Austria
Fetal Bovine Serum (FBS)	Gibco Invitrogen, Paisley, Scotland, UK

Table 7: Supplements and manufacturing companies.

Supplements	Company
Dimethylsulfoxide (DMSO)	Sigma-Aldrich, St. Louis, USA
Ethylenediaminetetraacetic acid (EDTA)	Carl Roth GmbH, Karlsruhe, Germany
Trypsin	Carl Roth GmbH, Karlsruhe, Germany
Cordycepin	Sigma-Aldrich, St. Louis, USA

2.2 Methods

2.2.1 In silico

2.2.1.1 Forming the cohort

To extract unsolved autosomal recessive ataxia families for further analysis, we started with a database (February 2017) containing all patients presenting at the department of neurology at the University of Tübingen with symptoms of cerebellar ataxia or hereditary spastic spinal paralysis and an available exome and/or genome dataset.

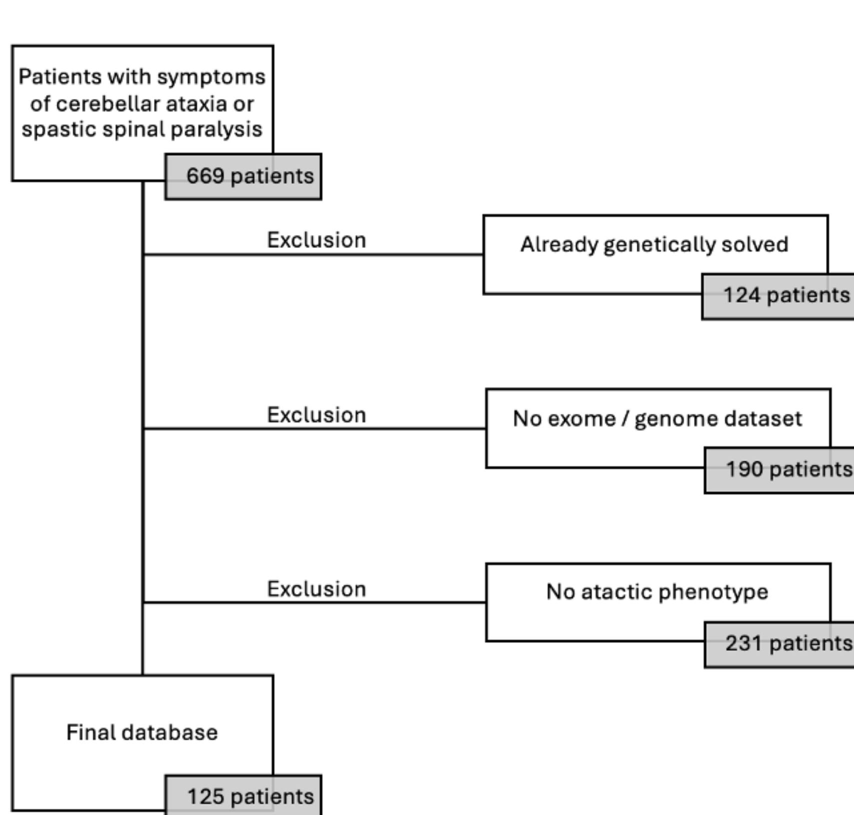

Figure 3: Starting from the original database of 669 patients, systematic sorting and exclusion led to our final database of 125 patients.

As shown in Figure 3 the initial table included 669 entries. First all patients that were already solved genetically were removed. Afterwards a systematic sorting according to the phenotypes presented by the remaining patients was performed. It was differentiated between patients that showed atactic symptoms and those with symptoms of a spastic spinal paralysis. There were 125 patients in 108 families left with family members

showing atactic signs like deficits in balance and coordination. There are also non-cerebellar features like a mental retardation and facial dysmorphisms.

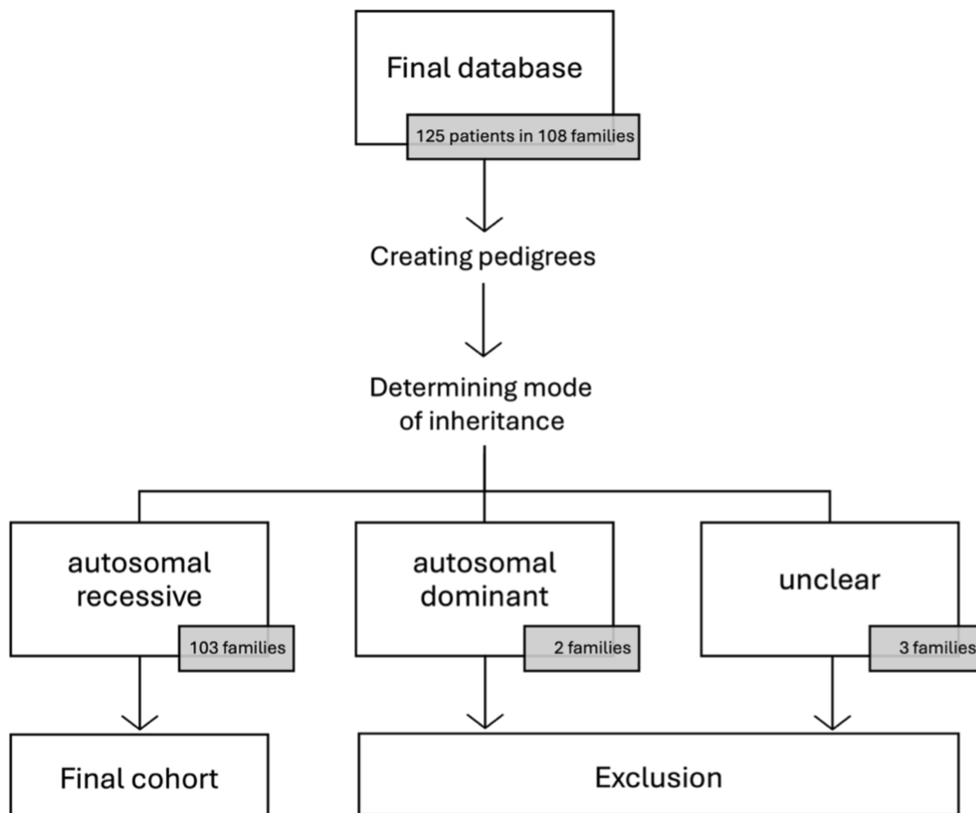

Figure 4: The final dataset was analysed to determine the most likely inheritance pattern for each family. Ultimately, 103 families were selected to form the final cohort.

As shown in Figure 4 a pedigree was then constructed for each family and the most likely mode of inheritance was determined. Three families were excluded from the final cohort, because no mode of inheritance could be determined due to insufficient information on the family members. Two further families were also excluded from the final cohort, because the most likely mode of inheritance was suspected to be autosomal dominant. In this study it was focused on the families most likely to have an autosomal recessive inheritance.

After extensive sorting, 103 families finally remained to form our final cohort.

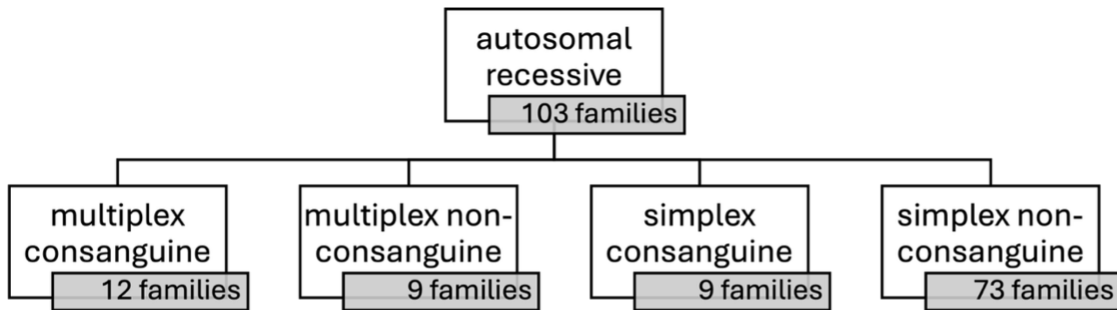

Figure 5: The resulting final cohort consisting of 103 families was subdivided into simplex and multiplex cases and with regard to consanguinity.

The final cohort is further subdivided as shown in Figure 5. It was divided into simplex and multiplex cases. A simplex case means that only one individual in the family is affected, whereas a multiplex case means that several individuals are affected. Consanguinity was also considered, and simplex or multiplex cases were classified accordingly.

2.2.1.2 NGS data analysis using the Genesis platform

Genesis is a unique, cloud-based, NIH-supported online database to analyse exome and genome data sets. These datasets are uploaded by different users and are subsequently available to all so that data alignment can take place to discover new disease-causing variants and moreover to also confirm new candidate genes and their characterization. In 2015 already 12.403 whole exome / whole genome datasets have been uploaded with more than 100 defined phenotypes.

Several individual filter strategies are possible and can also combine phenotype and genotype. The interpretation of the results is then carried out by the expert by performing further investigations of the variant found. (Gonzalez *et al.*, 2015)

The following filter criteria were used in this study:

1. Allele frequencies

   Allele frequency is defined as the relative frequency of a variant within a population. It is calculated by dividing the number of copies of a particular allele by the total number of copies of all observed alleles in that locus in a population. (Gillespie, 2004)

2. Quality filters

   Quality filters are divided into depth filters and genotype quality filters.

   The depth filter is necessary for achieving more accurate genotype results as it describes the number of reads that passed quality control. The higher the reads are, the better the genotype results. The minimum value should be eights reads to have a less than 1% chance of a biallelic variant appearing monoallelic.

   The genotype quality corresponds to the Phred quality score and indicates the probability that a specific genotype is the true genotype. Similar to the depth filter, a higher score means a more accurate result. Its minimum value is 20 with a corresponding Phred quality score of 99% accuracy. (Carson *et al.*, 2014)

3. Genesis allele counts

   The Genesis allele count restricts the maximum number of variants occurring in the Genesis database. The higher this value is, the more individuals show the same allele variant.

4. Inheritance pattern

   In this study it was differentiated between following inheritance patterns:

   - De novo
     - simplex cases without other affected family members and one possible disease-causing variant
   - Autosomal dominant
     - multiplex cases with affected members in every generation
   - Homozygous recessive
     - simplex-consanguineous cases with healthy individuals and an affected child
     - multiplex cases with affected individuals in not every generation
   - Compound heterozygous recessive
     - simplex or multiplex cases with healthy individuals and affected children

5. Variant Class

   It was differentiated between loss of function, non-synonymous and synonymous variants.

A loss of function variant leads to a change in the transcript, which in turn leads to a direct change in the protein. This group includes transcript ablation, a new splice acceptor or donor, stop gaining variants, a frameshift or loss of start or stop codon. These variants are disease-causing only in recessive inheritance or in the presence of haploinsufficiency. (Houge *et al.*, 2021)

Non-synonymous in contrast to synonymous variants alter the DNA- and subsequently the amino acid sequence, possibly being disease-causing.

Synonymous variants are thought to be neutral because the resulting protein does not show any change in its amino acid sequence. However there are cases in which a synonymous variant appears to be disease causing due to alterations in splicing sites. (Akli *et al.*, 1990; Richard and Beckmann, 1995; Griffiths, 2002; Pagani, Raponi and Baralle, 2005)

6. Genomic locus

   When searching for variants in already known ataxia-causing genes as well as searching for second hits of candidate genes, there is also the possibly to filter for a specific genomic locus.

2.2.1.2.1 Analysis of variants in known ataxia genes

First, all affected individuals in our cohort were analysed for previously known ataxia-causing gene variants.

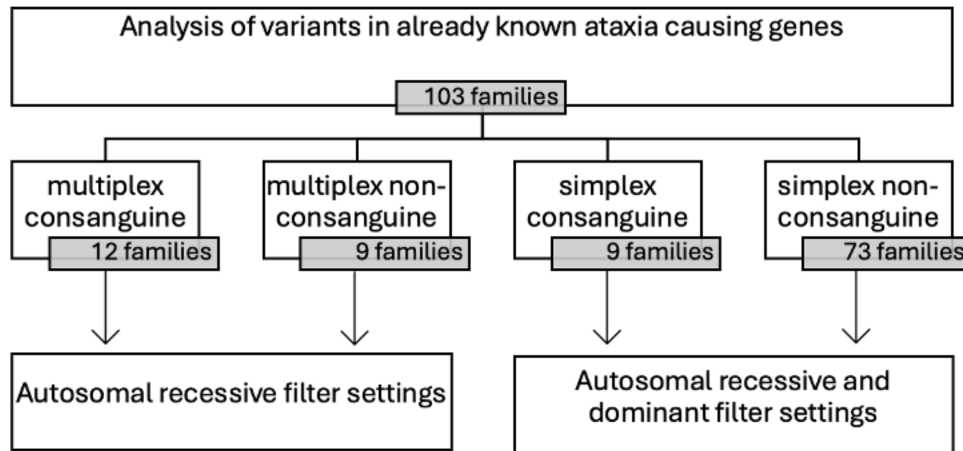

Figure 6: All 103 families forming the final cohort were first analysed for the presence of variants in already known ataxia-causing genes.

As shown in Figure 6 family by family analysis was performed in Genesis for already known variants shown in Table 8. We distinguished between dominant and recessive inheritance, since in general both would be possible. The probability of a recessive inheritance pattern is higher in simplex cases with consanguineous family background, but also a dominant inheritance pattern or de novo variant could be possible. In multiplex cases with affected individuals but healthy parents a recessive inheritance pattern is assumed.

The following filter settings were applied: quality filters DP > 8, GQ > 50, QUAL > 35, genomic locus of all known ataxia causing genes.

For an autosomal dominant inheritance, the minor allele frequency in ExAC, 1000 Genomes and EVS was set to be <0,01% and the Genesis allele count had been set to 5.

For an autosomal recessive inheritance, the minor allele frequency in ExAC, 1000 Genomes and EVS was set to be <1% and the Genesis allele count had been set to 50.

Table 8: Known ataxia genes (table provided by Prof. Dr. med. R. Schüle-Freyer via the European Reference Network for Rare Neurological Diseases (ERN RND), status as of 05/2017).

A

ABCB7, ABCD1, ABHD12, ADGRG1, AFG3L2, AHI1, ALG6, AMACR, ANO10, APTX, ARL13B, ARL3, ARMC9, ARSA, ATCAY, ATM, ATN1, ATP1A3, ATP2B3, ATP7B, ATP8A2, ATXN1, ATXN10, ATXN2, ATXN3, ATXN7, ATXN8, ATXN8OS

B

B4GALNT1, B9D1, BEAN1, BTD

C

CA8, CACNA1A, CACNA1G, CACNB4, CAMTA1, CC2D2A, CCDC88C, CEP104, CEP120, CEP290, CEP41, CLCN2, CLN5, CLN6, CLN8, COQ2, COQ6, COQ8A, CP, CPLANE1, CSPP1, CTDP1, CWF19L1, CYP27A1, CYP7B1

D

DAB1, DARS2, DDB2, DLAT, DNAJC19, DNAJC3, DNAJC5, DNMT1

E

EEF2, EIF2B1, EIF2B2, EIF2B3, EIF2B4, EIF2B5, ELOVL4, ELOVL5, ERCC2, ERCC3, ERCC4, ERCC5

F

FGF14, FLVCR1, FMR1, FOLR1, FXN

G

GALC, GBA, GBA2, GCDH, GFAP, GJB1, GLB1, GLRX5, GOSR2, GRID2, GRM1

H

HEXA, HEXB, HPRT1, HSD17B4, HYLS1

I

IFRD1, INPP5E, INPP5K, ITPR1

K

KCNA1, KCNA2, KCNC1, KCNC3, KCND3, KCNJ10, KIAA0556, KIAA0586, KIAA0753, KIF1A, KIF1C, KIF7

L

L2HGDH

M

MARS2, MFSD8, MKS1, MLC1, MME, MRE11, MSTO1, MTPAP, MTTP

N

NDUFS7, NEU1, NKX2-1, NKX6-2, NOP56, NPC1, NPC2, NPHP1, NUBPL

O

OFD1, OPA1, OPA3

P

PCNA, PDE6D, PDHX, PDSS1, PDSS2, PDYN, PEX1, PEX10, PEX11B, PEX12, PEX13, PEX14, PEX16, PEX19, PEX2, PEX26, PEX3, PEX5, PEX6, PEX7, PHYH, PIBF1, PIK3R5, PLA2G6, PLEKHG4, PMM2, PMPCA, PNKP, PNPLA6, POLG, POLR3A, POLR3B, PPP2R2B, PPT1, PRICKLE1, PRKCG, PRPS1, PRRT2, PTRH2

R

RARS2, RNF168, RNF170, RNF216, RNU12, RPGRIP1L, RUBCN

S

SACS, SCA20, SCA25, SCA30, SCA32, SCA37, SCYL1, SEPSECS, SEPT9, SETX, SIL1, SLC17A5, SLC19A3, SLC1A3, SLC25A15, SLC25A46, SLC2A1, SLC52A2, SLC9A1, SLC9A6, SMPD1, SNAP25, SNX14, SPART, SPG7, SPR, SPTBN2, SQSTM1, STUB1, SYNE1, SYT14

T

TBCE, TBP, TCTN1, TCTN2, TCTN3, TDP1, TDP2, TGM6, TMEM138, TMEM216, TMEM231, TMEM237, TMEM240, TMEM67, TPP1, TRAPPC11, TRPC3, TSEN2, TSEN34, TSEN54, TTBK2, TTC19, TTPA, TUBB2B, TUBB4A, TWNK

V

VAMP1, VARS2, VLDLR, VPS13A, VRK1

W

WDR73, WDR81, WWOX

X

XPA, XPC

Z

ZFYVE26, ZNF423, ZNF592

2.2.1.2.2 Analysis of variants in novel candidate genes

The remaining families that could not be solved with variants in already known ataxia-causing genes were analyzed regarding new candidate genes.

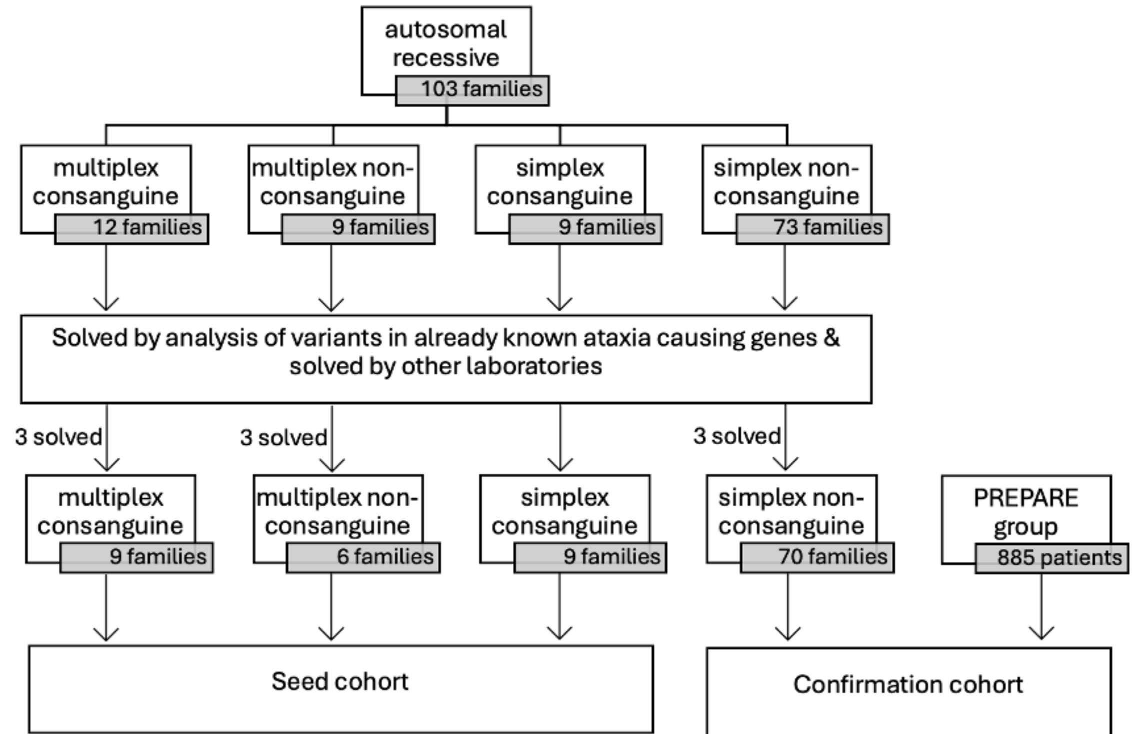

Figure 7: After exclusion of the families that were solved either by already known genes or by other laboratories, the remaining multiplex and simplex consanguine families formed the seed cohort. All simplex non-consanguine families and the PREPARE group formed the confirmation cohort.

The first step was to create the seed cohort consisting of all multiplex cases and simplex-consanguineous cases as shown in Figure 7. In total these are 24 families. Assuming an autosomal recessive inheritance matter, the filter settings were adapted to be stricter with the minor allele frequency in ExAC, 1000 Genomes and EVS being set to <0,2% and the Genesis allele count being set to 11. The results were reviewed regarding their segregation within the families. Afterwards a second hit search in our confirmation cohort was performed using segregating variants to identify other affected individuals. Our confirmation cohort consists not only of all our non-consanguineous simplex cases but includes also the ataxia datasets collected by the PREPARE consortium.

"PREPARE" is a research project that established a network of patients with recessive ataxias from 2016 to 2019 with the aim of developing systematic therapy studies including basic genetic characterisation, drosophila and mouse model and biomarker

screening. The PREPARE consortium has collected whole exome and genome sequencing datasets from unsolved ataxia families that are accessible via the Genesis database. In total it contained 885 patients in 825 families. We used the PREPARE dataset to validate candidate variants resulting from our approach in a larger cohort of cases.

The variants found were then further screened for potential pathogenicity.

2.2.1.3 Omim

To validate a possible candidate gene and to collect more details about specific genes the database https://www.omim.org/ was used.

2.2.1.4 ExAC and GnomAD

ExAC

The Exome Aggregation Consortium browser (ExAC browser) is large databank that contains exome sequence datasets from 60.706 individuals. (Karczewski *et al.*, 2017) These individuals are of diverse ethnicities. This browser contains nine times more exomes than the Exome Variant Server (EVS) and 24 times more Exomes than the 1000 Genomes (1000G) Project. It provides a publicly available platform for the evaluation and interpretation of clinical data and genetic information from affected individuals. (Lek *et al.*, 2016)

GnomAD

The Genome Aggregation Database (GnomAD v4.0) consists of 125.748 exomes and 15.708 genomes. It has replaced the ExAC browser and contains all its data. (Karczewski *et al.*, 2020)

2.2.1.5 Constraint scores

The intolerance of a gene regarding a specific type of variant is indicated by means of the constraint score. The fewer variations of a particular class a gene shows, the more intolerant it is to that specific type of variation. A distinction is made between Z-scores

and pLI-scores. The Z-scores refer to the missense and synonymous categories, whereas the pLI-scores are used for the loss-of-function variations.

Positive Z-scores indicate an increased intolerance to a type of variation, so that the gene accordingly has fewer variants of this type. Negative Z-scores are linked to a higher tolerance towards a variant and therefore more variants can be found.

The probability of a specific variant being loss-of-function intolerant is refelected with the pLI-score. There are three possible classes regarding the genes' tolerance towards loss of function variants: null, recessive and haploinsufficient. Null means that it shows full tolerance. In a recessive inheritance manner the heterozygous loss of function variant is tolerated, whereas a homozygous variant is not. In case of haploinsufficiency there is no tolerance of a heterozygous loss of function variant. The pLI-score that is ≥0,9 indicates a high intolerance regarding a loss of function variant. (Samocha *et al.*, 2014, 2017; Lek *et al.*, 2016)

2.2.1.6 Conservation scores

As species develop changes in the genome might occur. During evolution different species underwent different changes that could either be tolerated or be pathogenic. Whole exome and genome sequencing has been used to trace evolutionary changes in DNA base sequences. Comparing the exomes and genomes of different species, similarities and differences can be identified. In particular, the sections that have persisted across the different species are considered to be highly conserved and therefore rather intolerant of changes, so that changes here have an increased pathogenic potential. Typically, natural selection eliminates variants and variants in regions that are highly conserved.

PhastCons and PhyloP are two methods to measure the evolutionary conservation.

PhastCons is a method based on a hidden Markov model. Using multiple alignments, it estimates each nucleotides possibility whether it belongs to a conserved element. It does not only consider each alignment column by its own but also pays attention to neighbouring columns and shows the probabilities of a negative selection. The PhastCons score ranges between 0 and 1 with values closer to 1 indicating higher conservation.

In contrast, PhyloP is a method that only measures each individual columns conservation separately, ignoring the effects of flanking columns. It also is a measurement for both acceleration and conservation. Positive scores indicate well conserved sites, negative scores are interpreted as fast evolving. (Yang, 1995; Siepel *et al.*, 2005; Pollard *et al.*, 2010)

2.2.1.7 CADD Score

The CADD score estimates the probability that a genetic variant, such as a single nucleotide or insertion/deletion variant, will have a harmful effect on the human genome. To this end, it combines several annotations, including conservation and function information, into a single metric. Variants with higher scores are more likely to be harmful. (Rentzsch *et al.*, 2019)

2.2.1.8 ACMG

To evaluate variants, the American College of Medical Genetics and Genomics together with the Association for Molecular Pathology and the College of American Pathologists developed a standardised classification system. There are five categories: benign, likely benign, uncertain significance, likely pathogenic and pathogenic. (Richards *et al.*, 2015)

This classification can be derived into a points system that is based on the Bayesian principle. This results in the scores shown in Table 9, which are proportional to the logarithmic odds ratio. (Tavtigian *et al.*, 2020)

Table 9: ACMG classification point system to interpret variants (source: Tavtigian et al., 2020).

Category	Point ranges
Pathogenic	≥ 10
Likely Pathogenic	6 - 9
Uncertain	0 - 5
Likely Benign	-1 - -6
Benign	≤ -7

The ACMG classifications were taken from the VarSome germline classification from the database https://varsome.com (version 11.4).

2.2.1.9 Tissue expression

A change at the DNA level does not necessarily translate into a change at the protein level that ultimately results. Since the hereditary ataxias are particularly characterized by alteration or loss of neuronal tissue in the brain and cerebellum, a look at the tissue expression of the candidate genes and their transcripts is also important.

To be able to classify the functional effects Genotype-Tissue-Expression project (GTEx) and Human Protein Atlas portal were used as sources to have a look particularly at expression in the brain and cerebellum.

The GTEx project (https://gtexportal.org/home/) is a database containing 17.382 samples of 948 donors (as of 2021). These donors are of different sex and ethnical and racial groups. They are between 21 - 70 years old and sample collection started within 24 hours of their deaths. Exclusion criteria are an infection with the human immunodeficiency virus or hepatitis, cancer in metastatic situation, chemotherapy or radiation within the previous two years, blood transfusions within 48 hours prior to the death or the body mass index below 18.5kg/m² or above 35kg/m². (Lonsdale *et al.*, 2013; 'The Genotype-Tissue Expression (GTEx) project', 2013)

The Human Protein Atlas portal (https://www.proteinatlas.org/) contains samples from 144 cancer patients. These samples were taken from normal tissue around the diseased area and revised by a pathologist and found to be histologically healthy. (Pontén, Jirström and Uhlen, 2008; Uhlén *et al.*, 2015)

2.2.1.10 Interaction between known ataxia genes and novel candidate genes

The interaction of our candidate genes with already known ataxia-causing genes was investigated using the high-throughput identification pipeline for promoter interacting enhancer elements "HIPPIE" (http://cbdm-01.zdv.uni-mainz.de/~mschaefer/hippie/). This is a pipeline using Hi-C raw reads to identify DNA-DNA interacting fragments. (Hwang *et al.*, 2015)

We examined both the direct interaction and the interaction between genes that occurs via a maximum of 1 additional gene.

2.2.1.11 Primer design

To find the perfect fitting primers for each individual genetic position there were 4 steps necessary. First the genomic position of each variant was given by Genesis. With that data the UCSC Genome Browser (http://genome.ucsc.edu/index.html) was then used to get the exact DNA region adding 200 bases up- (5') and downstream (3'). The DNA region was copied and pasted to Primer3web (http://bioinfo.ut.ee/primer3/), the variant was put in parentheses and the program put out several left and right primers. These primers where then checked using the primer blast tool by NCBI (https://www.ncbi.nlm.nih.gov/tools/primer-blast/). Criteria for the primer to be suitable include a low self-3' complementarity and less than three mismatches.

2.2.2 In vitro

2.2.2.1 Polymerase chain reaction

The polymerase chain reaction (PCR) is a well-established procedure in the laboratory to amplify several DNA fragments. First specific primers were designed (as described in 2.2.1.2) matching to the DNA region of interest. Following this the DNA amplification was performed using the reaction shown in Table 10 and the standard PCR protocol shown in Table 11.

Table 10: PCR reaction mix.

Reagent	Amount (µl)
Green GoTaq Reaction Buffer 5x	5
dNTPs (10 mM)	0.5
G2 GoTaq DNA Polymerase	0.25
Primer (forward & reverse)	1.25
DNA sample (10ng/µl)	2.5
ddH_2O	15.5
Total	25

Table 11: PCR protocol.

Step	Temperature (°C)	Duration
1	95	5 min
2	94	30 sec
3	62	30 sec
4	72	30 sec
5		Go to step 2; repeat 35x
6	72	10 min
7	4	forever

2.2.2.2 Touchdown PCR

Touchdown polymerase chain reaction is used to increase the specificity of the reaction. As shown in Table 12 the initial annealing temperature was at 62°C and gradually decreased by 1°C until the calculated annealing time for the designed primers was reached. The PCR was then carried out with this temperature.

Table 12: tdPCR protocol.

Step	Temperature (°C)	Duration
1	94	5 min
2	94	30 sec
3	62 (-1°C/cycle)	30 sec
4	72	30 sec
5		Go to step 2; repeat 10x
6	94	30 sec
7	52	30 sec
8	72	30 sec
9		Go to step 6; repeat 25x
10	72	7 min
11	4	forever

2.2.2.3 Gel Electrophoresis

To confirm whether the PCR worked and the correct product was amplified, a gel electrophoresis was carried out for each PCR approach.

For this purpose, a 2% agarose gel with the components listed in Table 13 is prepared. First agarose and TBE buffer were mixed and heated in the microwave until a uniform gel was obtained. Then Midori Green was added, stirred and the gel was poured in an appropriate mold, everything cooled down and formed a solid gel.

The first chamber of the 2% agarose gel was loaded with 7µl gene ruler DNA ladder mix as size marker, the other chambers were loaded with 5µl of the DNA sample.

A voltage of 200V for large gels respectively 120V for small gels was applied for 30 minutes. Afterwards the gel was evaluated via fluorescence.

Table 13: Components 2% agarose gel.

Component	Amount
SeaKem LE Agarose	4g
TBE buffer	200ml
Midori Green	12µl

The remaining PCR products were purified with sodium acetate and ethanol (NaAc/EtOH). Therefore, 40µl NaAc/EtOH were added to 20µl of the remaining PCR product and centrifuged at 23°c for 45 minutes with 3220 rcf. The supernatant was discarded and as a washing process 100µl of 70% ethanol was added and centrifuged at 23°C for 10 minutes with 3220 rcf. The supernatant was discarded again and the washing process was repeated once. As last purification step the tubes were centrifuged upside down without their caps at 23°C with 600 rcf for 1 minute.

For further use the now precipitated PCR products were diluted in 15µl ddH_2O and gently mixed for 30 minutes at 23°C.

2.2.2.4 Sequencing PCR

A sequencing PCR was performed to prepare the PCR products for the following Sanger Sequencing. The PCR was performed as described in section 2.2.2.1, with the exception that forward and reverse primers were used separately. The sequencing PCR reaction mix is shown in Table 14 and the sequencing PCR protocol is shown in Table 15.

Table 14: Sequencing PCR reaction mix.

Reagent	Amount (µl)
Colorless GoTaq Reaction Buffer 5x	1.65
BigDye™ v3.1	0.7
Primer (10pmol/µl) forward or reverse	1
Purified PCR product	4
ddH_2O	2.65
Total	10

Table 15: Sequencing PCR protocol.

Step	Temperature (°C)	Duration
1	94	1 min
2	94	10 sec
3	50	5 sec
4	60	4 min
5		Go to step 2; repeat 30x
6	4	forever

The PCR products were purified with 30ml NaAc/EtOH. The further procedure remained the same as described in section 2.2.2.3.

2.2.2.5 Sanger Sequencing

To determine the specific nucleotide sequence of the PCR products Sanger sequencing was used. The wells of a sequencing plate were loaded with 15µl of Formamide and 7µl

of the purified PCR products where added. Subsequently the samples were sequenced with the 3100 Avant Genetic Analyzer using the standard instrument protocol. The results were afterwards analysed with Pregap4.

2.2.2.6 Fibroblast culture and RNA Isolation

To isolate the RNA from fibroblasts the patient derived cells were thawn in a water bath, afterwards treated with DMEM and FBS and centrifuged to clear the cells. The supernatant was removed and the resulting pellet was resuspended with 100µl of medium. The cell suspension was then added into a $25cm^3$ flask with 3ml DMEM and FBS. After two days the medium changed and the cells were fed. They were splitted after four days using trypsin and EDTA to detach the cells from the flask. The reaction was stopped with DMEM and the cells were transferred into new and bigger flasks. To isolate the RNA from those fibroblasts the High Pure RNA Isolation Kit from Roche was used.

2.2.2.7 Cordycepin

To check on the polyadenylation status we treated the above-mentioned fibroblasts of our control group with Cordycepin 50µl, because Cordycepin is known to prevent the polyadenylation of the mRNA. The fibroblasts of our control group were divided into three parts. The first part was treated with 50µl Cordycepin for 15 minutes, the second part was treated with 50µl Cordycepin for 45 minutes and the third part was not treated with Cordycepin. Our aim was to compare the length of the polyadenylation of the untreated and treated control with our patients' polyadenylation length.

2.2.2.8 Polyadenylation Assay

To determine the length of poly-a-tails the USB®-Poly(A) Tail-Length Assay Kit from Affymetrix was used. Fibroblasts were cultured and RNA was isolated as described above.

First poly(G/I)-tails were added by adding the reagents shown in Table 16 to the RNA sample.

Table 16: Poly(G/I)-tail-mix used in the polyadenylation assay.

Reagent	Amount (µl)	Final concentration
Total RNAsample, 1µg	Up to 14µl	0,1µg up to 2µg
5x Tail Buffer Mix	4µl	
10x Tail Enzyme Mix	2µl	
Water, nuclease-free	Up to 20µl	
Total	20µl	

Afterwards this mixture was incubated for 60 minutes at 37°C. 2µl of 10x Tail Stop Solution were added and the now poly(G/I)-tailed RNA was reverse transcribed by following the protocol described in Table 17.

Table 17: Reverse transcription mix used in the polyadenylation assay.

Reagent	RT +	RT -
G/I Tailed RNA Sample	5µl	5µl
5x RT Buffer Mix	4µl	4µl
10x RT Enzyme Mix	2µl	-
Water, nuclease-free	9µl	11µl

The resulting probes were incubated at 44°C for 60 minutes, next at 92°C for 10 minutes. The amplification was performed directly after the incubation time. Therefore, the samples were diluted by adding 20µl nuclease-free water, added to the PCR mix shown in Table 18 and the PCR was performed using the protocol described in Table 19.

Table 18: PCR mix used in the polyadenylation assay.

Reagent	RT + tail PCR	RT – tail PCR	RT + specific PCR	RT – specific PCR
Diluted RT sample	2,5µl	2,5µl	2,5µl	2,5µl
5x PCR Buffer Mix	5µl	5µl	5µl	5µl
10µM Gene-Specific PCR Forward Primer	1µl	1µl	1µl	1µl
10µM Universal PCR Reverse Primer	1µl	1µl	-	-
10µM Gene-Specific PCR Reverse Primer	-	-	1µl	1µl
25mM MgCl2	1,5µl	1,5µl	1,5µl	1,5µl
1.25 units/µl HotStart-IT® Taq DNA Polymerase	1µl	1µl	1µl	1µl
Water, Nuclease-Free	13µl	13µl	13µl	13µl

Table 19: PCR protocol used in the polyadenylation assay.

Step	Temperature (°C)	Duration
Pre-incubation	94	2 min
Amplification	94	10 sec
	60	30 sec
	72	5 min
Soak	4	forever

The detection was performed using two different methods. First the size of the PCR product was assessed with a 2% agarose gel and visualized with a fluorescence image scanner. Second the PCR products were sequenced using the standard sequencing protocol described in 2.2.2.4 and in Tables 14 and 15.

3 Results

3.1 Overview of cohort

The aim of this work was to identify autosomal recessive ataxia genes. The focus is set on either homozygous or heterozygous recessive inheritance pattern.

As previously described our cohort was formed of patients who show ataxic symptoms with recessive inheritance being most likely. Our final cohort consists of 103 families who were not yet diagnosed genetically.

These 103 families were now further subdivided into simplex and multiplex cases as well as consanguineous and non-consanguineous families.

A simplex case is defined with only one affected individual in the entire family, whereas a multiplex case includes multiple affected individuals.

Consanguineous means biologically related. In consanguineous relationships, there is a higher risk of transmitting recessively inherited diseases as there is a higher probability that two individuals are carrying the same recessive variant.

This results in the following subdivision of the index cases, shown in Figure 8:

- 73 simplex non-consanguineous cases,
- 9 simplex consanguineous cases,
- 9 multiplex non-consanguineous cases and
- 12 multiplex consanguineous cases.

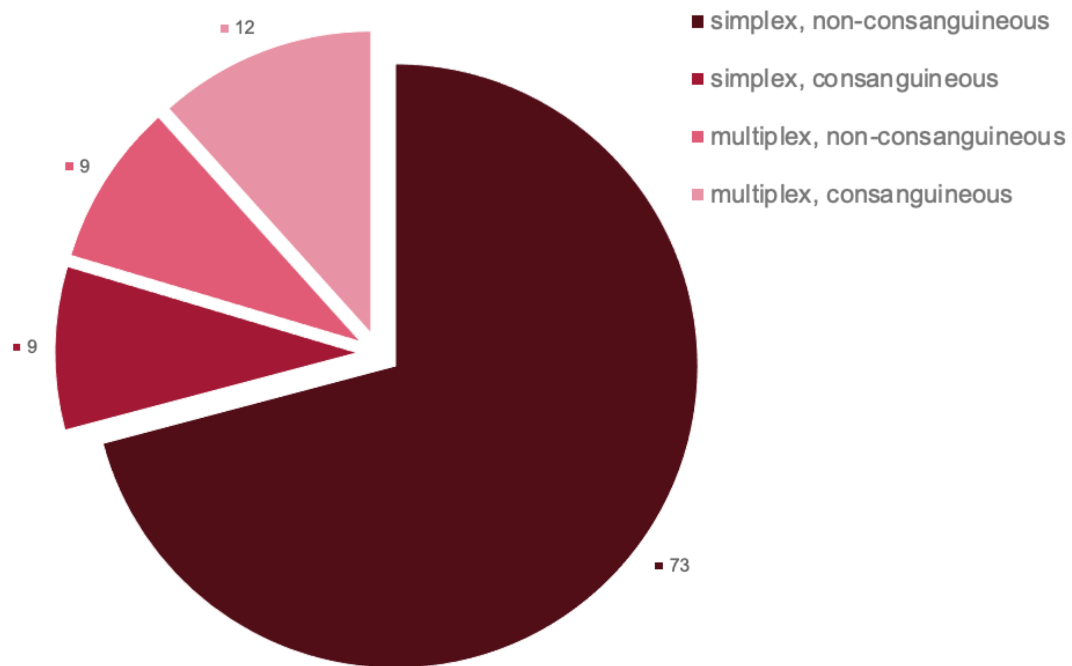

Figure 8: Distribution of our 103 index cases into simplex and multiplex cases as well as consanguineous and non-consanguineous cases.

In addition to the index cases, our cohort also includes other affected family members of the multiplex families as well as healthy relatives. All in all there are:

- 103 index cases,
- 13 other affected and
- 10 healthy family members.

Furthermore, we have whole exome and / or whole genome datasets for the individuals in our cohort. Our cohort includes:

- 68 individuals with whole exome datasets,
- 40 individuals with whole genome datasets,
- 14 individuals with both exome and genome datasets and
- 4 individuals without datasets.

All cases were divided into two tiers. Tier 1 is the seed cohort that includes all consanguineous and multiplex recessive cases and is used to search for new genes. Tier 2 formed our confirmation cohort and includes all non-consanguineous simplex cases and additionally the ataxia datasets collected by the PREPARE consortium.

3.2 Known genes

To address our objective to identify variants in known ataxia genes we performed a systematic re-analysis of the available NGS datasets. Hereby, we queried the data for the presence of variants in known ataxia causing genes. According to the ERN-RND genelist, there are 251 genes that are known to cause hereditary ataxias linked to autosomal dominant, recessive, X-linked or mitochondrial inheritance patterns.

In this study we only distinguished between autosomal dominant and recessive inheritance. Filter criteria that support the autosomal recessive or dominant inheritance pattern were applied (see methods) as well as the known ataxia causing genes (Table 8).

Each of the 103 families was examined in Genesis.

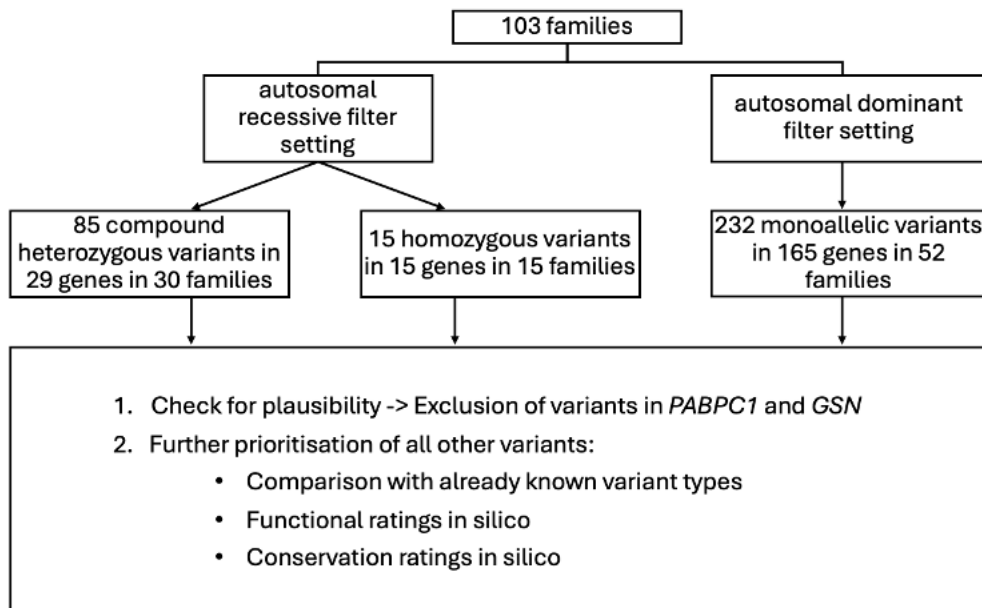

Figure 9: Procedure for searching in already known ataxia-causing genes. All variants found are checked for plausibility and then prioritised further.

In 67 families variants in already known ataxia-causing genes were identified in silico.

As shown in Figure 9 it was further differentiated between homozygous and compound heterozygous variants using the filter mode that supports autosomal recessive inheritance.

In simplex non-consanguine cases filter settings that support the autosomal dominant inheritance were also applied as an autosomal dominant mode of inheritance or de novo variant could also be possible.

These hits were checked in silico for plausibility. In the two genes *PABPC1* and *GSN*, a comparatively large number of variants was found that were therefore not investigated further, as the probability that these are pathogenic with so many occurring variants is unlikely. In the *PABPC1* gene 16 compound heterozygous variants in 9 families and in the *GSN* gene 10 hits in compound heterozygous variants in 1 family were found.

Subsequently, all other variants were further prioritised to identify truly potential disease-causing variants. All hits were compared with already known variant types that cause ataxias and especially truncating variants were further examined. Furthermore, the functional and conservation ratings of each variant found in silico were evaluated. The higher the functional and conservation score, the higher the probability that the variant found is potentially pathogenic.

First, segregation within families was examined in silico. Those variants that segregated in silico were further examined via PCR and Sanger sequencing.

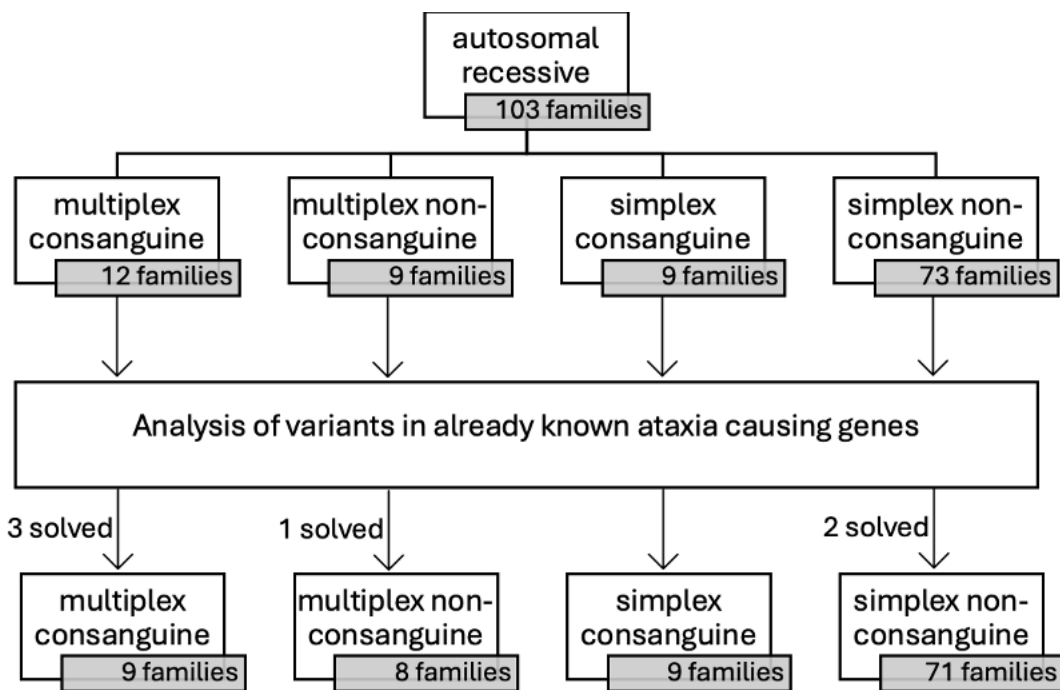

Figure 10: This is an overview of the number of families that could be genetically solved by analysing known genes.

With this approach 6 families could be genetically diagnosed as shown in Figure 10.

Table 20: Overview of the segregating variants in known ataxia genes.

Family	Chromosome	Position	Reference	Alternate	Genotype	Gene	Variant type
16	10	30629278	G	A	G/A	*MTPAP*	Missense
29	18	12340322	G	T	G/T	*AFG3L2*	Missense
29	18	12351346	G	A	G/A	*AFG3L2*	Missense
53	17	57775069	CAG	C	deletion / deletion	*PTRH2*	Frame shift
53	17	57775279	C	T	T/T	*PTRH2*	Missense
58	7	5752394	C	T	C/T	*RNF216*	Missense
58	7	5780868	CA	C	wildtype / deletion	*RNF216*	Deletion
59	22	50885968	C	G	C/G	*SBF1*	Missense
59	22	50898096	C	A	C/A	*SBF1*	Missense
81	4	25125738	C	T	T/T	*SEPSECS*	Missense
88	3	25770698	TTTGA	T	deletion / deletion	*NGLY1*	Deletion

Table 20 summarises the variants that we found in genes that are already known to cause ataxia. We also identified a segregating variant in the already known ataxia-gene *MTPAP*, but functional analysis showed no harmful impact on protein level.

3.2.1 *AFG3L2* – Family 29

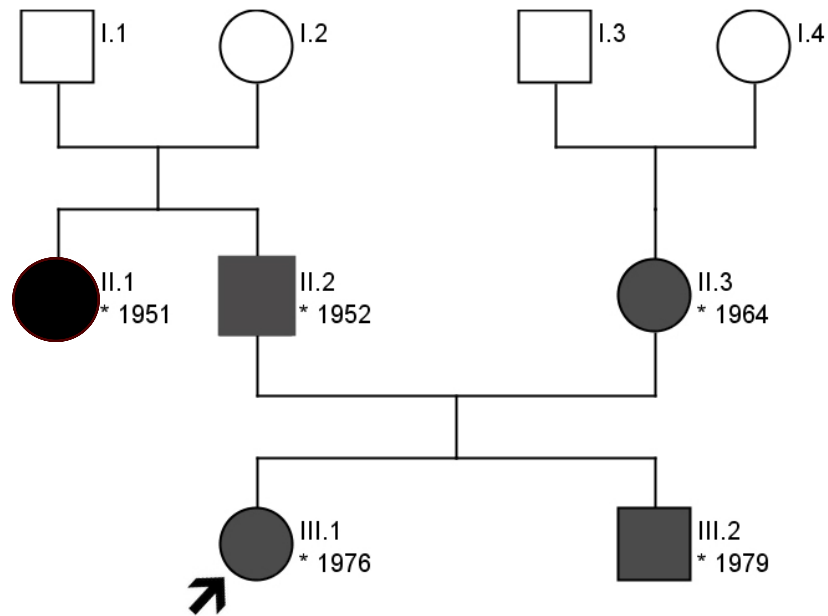

Figure 11: Pedigree of family 29.

Table 21: Overview of the AFG3L2 variants.

Chromosome	18	18
Position	12340322	12351346
Variant type	SNV	SNV
Reference	G	G
Alternate Allele	T	A
Zygosity	heterozygous	heterozygous
Gene	*AFG3L2*	*AFG3L2*
Genbank transcript ID	NM_006796.3	NM_006796.3
cDNA	c.1858C>A	c.1385C>T
Protein effect	p.Q620K	p.A462V
GnomAD allele frequency	0.000001590	0.000002030
ClinVar status	Pathogenic	Uncertain significance
ACMG-Classification	Pathogenic: 10 points = 10 P – 0 B	Pathogenic: 15 points = 15 P – 0 B
PhastCons100way	1.000	1.000

PhyloP100way	9.818	9.783
Index III.1	G/T	G/A
Mother II.3 (affected)	G/G	G/A
Brother III.2 (affected)	G/G	G/A
Grandmother I.2 (healthy)	G/G	G/G
Aunt II.1 (affected)	G/T	G/G
Father II.2 (affected)	G/T	G/G

In family 29 two heterozygous missense variants in the *AFG3L2* gene described in Table 21 were found with whole-genome sequencing and their segregation within the family was already confirmed via PCR and subsequent Sanger sequencing prior to this study. Interestingly, only the index patient carries both missense variants, whereas only one variant could be detected in each of the other affected family members.

First there is a heterozygous missense variant at the genomic position chr18:12340322 leading to a switch from guanine to thymidine with the new base triplet encoding for the amino acid lysine instead of glutamine. This is a site that is highly conserved (PhyloP100way 9.818 and PhastCons100way 1.000) and therefore being more intolerant towards variants.

The second heterozygous missense variant is located at the genomic position chr18: 12351346 with a switch from guanine to adenine. This new base triplet now encodes for the amino acid valine instead of alanine. Like the first variant this site is also highly conserved (PhyloP100way 9.783 and PhastCons100way 1.000). Both missense variants are classified as being pathogenic and the allele frequencies in the gnomAD database show that both variants are very rare. The z-Score of missense variants in the *AFG3L2* gene is 1,99 therefore showing less tolerance towards missense variants.

As shown in Figure 11 there are three family members showing signs of a neurological disorder and optic atrophy - the index patient, her mother and her younger brother. Two other relatives, the father and the paternal aunt, showed atactic symptoms in old age, but without optic atrophy.

The index patient presented with neurological symptoms since the age of 2. Initially she showed optic atrophy that progressed over time and lead to a bilateral blindness at age 35 years. At the age of 8 years a childhood myoclonus and at the age of 15 a spastic-atactic gait disorder was noticed. The ataxia was progressive and at the age of 35 she had a mild-to-moderate cerebellar ataxia with SARA score of 7 points, increasing to 13.5 when she was 43 years old. Furthermore, she presented with paraspasticity and a distal chorea of both hands.

Her mother reported early signs of a mild visual reduction starting at the age of 8 years that did not progress. She has a cervical dystonia without progression since being 55 years of age. She did not show any other neurological symptoms like a spastic-atactic gait disorder, myoclonus or chorea.

Her younger brother stated that since the age of 4 he had a progressive visual reduction. Like his mother he did not show other neurological symptoms.

Her father developed mild atactic symptoms at the age of 64 years with a SARA score of 4. He had no visual impairment or other neurological symptoms.

Interestingly, when comparing the different severity of neurological symptoms within the family, it quickly becomes apparent that the index patient is the most severely affected. The mild phenotype of the family members is due to the fact that they are each only carriers of a heterozygous variant. The index patient, on the other hand, shows a severe phenotype and carries bi-allelic variants.

3.2.2 *PTRH2* – Family 53

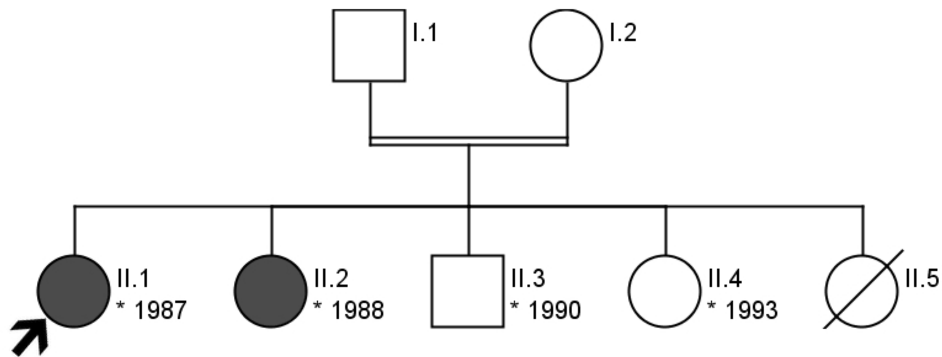

Figure 12: Pedigree of family 53.

Table 22: Overview of the PTRH2 variant.

Chromosome	17	17
Position	57775069	57775279
Variant type	INDEL	SNV
Reference	CAG	C
Alternate Allele	C	T
Zygosity	homozygous	homozygous
Gene	*PTRH2*	*PTRH2*
Genbank transcript ID	NM_016077.5	NM_016077.5
cDNA	c.269_270delCT	c.61G>A
Protein effect	p.Ala90GlyfsTer13	p.V21I
GnomAD allele frequency	Not present.	Not present.
ClinVar status	Pathogenic	Likely benign
ACMG-Classification	Pathogenic: 13 points = 13 P – 0 B	Benign: -17 points = 0 P – 17 B
PhastCons100way	1.000	0.999
PhyloP100way	9.675	3.689
Index II.1	deletion/deletion	T/T
Sister II.2 (affected)	deletion/deletion	T/T
Brother II.3 (healthy)	wildtype/deletion	C/C
Sister II.4 (healthy)	wildtype/deletion	C/T

In family 53 two variants in the *PTRH2* gene were found in this family with whole-genome sequencing and their segregation was confirmed via PCR and subsequent Sanger sequencing. As described in Table 22 there is a truncating frameshift variant at the genomic position chr17:57775069 leading to a deletion of the bases adenine and guanine, resulting in a premature stop codon and thus altering the amino acid sequence. It shortens the protein by 78 amino acids. This variant was not present in the gnomAD database. Its ACMG classification hints towards being pathogenic which is also supported by the fact that this is a homozygous truncating variant. This site also shows high conservation (PhastCons100way 1.000, PhyloP100way 9.675).

The second variant is a missense variant at the genomic position chr17:57775279 that leads to a switch from cytosine to thymine with the new base triplet encoding for the amino acid valine instead of isoleucine. This variant is classified as likely benign by the ACMG classification. It was not present in the gnomAD database and shows a high conservation (PhastCons100way 0.999, PhyloP100way 3.689), but slightly less than the first found variant.

In this family two out of five children born of consanguineous parents have an early onset neurological disorder as shown in Figure 12. The index patient II.1 as well as her sister II.2 suffer from movement disorders since their birth. Further symptoms are a severe hearing loss, mental retardation, hepatomegaly and steatosis hepatis, anaemia and a severe gait and balance impairment. They also suffer from vertigo and fatigue, stomach-ache and diarrhoea. Their condition does not seem to be progressive. There are three more children who appear healthy. One of those three sibling died of an unrelated cause.

3.2.3 *RNF216* – Family 58

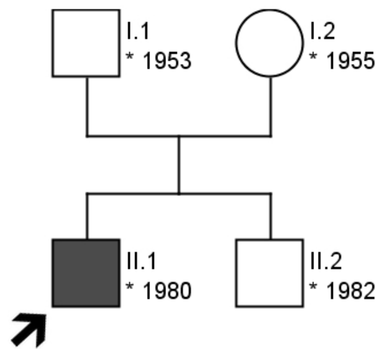

Figure 13: Pedigree of family 58.

Table 23: Overview of the RNF216 variant.

Chromosome	7	7
Position	5752394	5780868
Variant type	SNV	Deletion
Reference	C	CA
Alternate Allele	T	C
Zygosity	heterozygous	heterozygous
Gene	*RNF216*	*RNF216*
Genbank transcript ID	NM_207116.3	NM_207116.3
cDNA	c.1763G>A	c.608_608delT
Protein effect	p.Arg588Gln	p.Leu203ArgfsTer37
GnomAD allele frequency	Not present.	Not present.
ClinVar status	No data available.	No data available.
ACMG-Classification	Uncertain significance: 3 points = 3 P - 0 B	Likely pathogenic: 9 points = 9 P – 0 B
PhastCons100way	1.000	0.297
PhyloP100way	7.080	1.591
Index II.1	C/T	wildtype/deletion
Mother I.2 (healthy)	C/C	wildtype/deletion
Father I.1 (healthy)	C/T	wildtype/ wildtype
Brother II.2 (healthy)	C/C	wildtype/deletion

In family 58 we found two heterozygous variants in the *RNF216* gene, described in Table 23, segregating within the family with the index patient being a carrier of both variants whereas his relatives only carry one variant each.

There is a heterozygous missense variant at the genomic position chr7:5752394, leading to a switch from cytosine to thymine. This switch causes the amino acid glutamine being incorporated instead of arginine. This variant was not present in gnomAD and is of uncertain significance. It is a highly conserved site (PhastCons100way 1.000, PhyloP100way 7.080).

The second variant is a deletion at the genomic position chr7:5780868, which creates a new premature termination codon. This variant was also not present in gnomAD and it is classified by the ACMG classification as likely pathogenic.

As shown in Figure 13 there is a male patient suffering from an early onset neurological disorder in this non consanguineous family. The family history is blank for movement disorders and seizures. The parents appear healthy and are not consanguineous. The younger brother also appears healthy.

The patient states that his seizure disorder has existed since childhood and that since 2003 he has had progressive symptoms and a suspected cognitive impairment of unclear origin. Magnetic resonance imaging showed an infratentorial emphasized cerebral atrophy as well as changes in the medullary layers.

After a difficult birth with a healthy infant, who learned to sit at eight months, to walk at 14 months and to speak his first words at 18 months, a seizure occurred for the first time at the age of about six years. He was subsequently placed on a phenobarbital and later on levetiracetam.

At the age of 23 an unusual and progressively worsening gait pattern occurred and an increasing clumsiness of both hands was noticed. The patient also has problems pronouncing words correctly. Double vision and dysphagia are not present. The patient appears childlike and mentally retarded.

The neurological examination findings show clear cerebellar symptoms. The finger-following-test shows a hypermetria on the left side. A mild tremor and dysmetria emphasized as well on the left side is noticeable in the finger-nose-test. There is a mild

bradydiadochokinesia on his left side. The knee-heel-test shows a tremor and dysmetria. His stand is broad-based. Rhomberg's standing test is possible, with eye closure clear tendency to fall. Single-leg and tandem stance no longer possible with progression of the disease. There is an ataxic gait pattern. Free walking is no longer possible, only with help of a rollator and even than he is very unsteady.

Laboratory chemistry showed a decreased arylsulphatase in the urine as an indication of a possible metrechromatic leucodystrophy. However, a demyelinating polyneuropathy could not be diagnosed.

A muscle biopsy showed mild chronic neurogenic muscle atrophy. The nerve biopsy showed evidence of a moderate neuropathy, but without electrophysiological evidence.

Furthermore, hypogonadotropic hypogonadism and secondary osteoporosis were diagnosed.

3.2.4 *SBF1* – Family 59

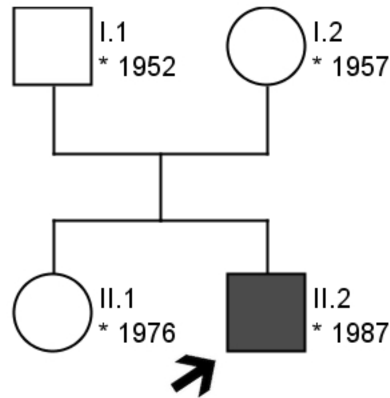

Figure 14: Pedigree of family 59.

Table 24: Overview of the SBF1 variant.

Chromosome	22	22
Position	50885968	50898096
Variant type	Missense	Splice acceptor variant
Reference	C	C
Alternate Allele	G	A
Zygosity	heterozygous	heterozygous
Gene	*SBF1*	*SBF1*
Genbank transcript ID	NM_002972.4	NM_002972.4
cDNA	c.5434G>C	c.3492-1G>T
Protein effect	p.Asp1812His	Non.
GnomAD allele frequency	Not present.	Not present.
ClinVar status	Uncertain significance	Likely pathogenic.
ACMG-Classification	Uncertain significance: 2 points = 2 P – 0 B	Pathogenic: 11 points = 11 P – 0 B
PhastCons100way	1.000	1.000
PhyloP100way	7.412	7.386
Index II.2	C/G	C/A

In family 59 using Genesis two variants in the *SBF1* gene were found, which are described in Table 24. First there is a missense variant at the genomic position Chr22:50885968 that leads to a change in the amino acid sequence by changing the aspartic acid to histidine at position 1812. This variant is classified with an uncertain significance. It is in a highly conserved location (PhastCons100way 1.000, PhyloP100way 7.412).

Second there is a truncating variant at the genomic position Chr22:50898096 that leads to a new splice acceptor in a non-coding area. This variant is classified as being pathogenic by the ACMG. Like the above-mentioned variant this variant is also highly conserved (PhastCons100way 1.000, PhyloP100way 7.386).

Both variants were not present in the gnomAD database.

In this non consanguineous family, as shown in Figure 14, there is a male patient who showed first atactic signs at the age of 3. His sister and parents appear healthy.

3.2.5 *SEPSECS* – Family 81

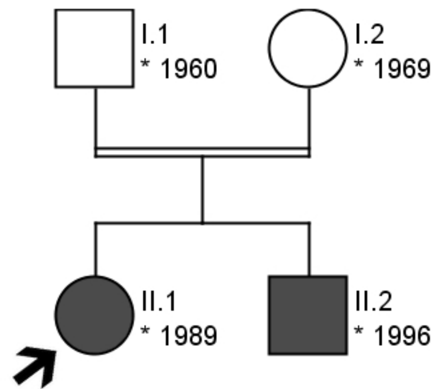

Figure 15: Pedigree of family 81.

Table 25: Overview of the SEPSECS variant.

Chromosome	4
Position	25125738
Variant type	SNV
Reference	C
Alternate Allele	T
Zygosity	homozygous
Gene	*SEPSECS*
Genbank transcript ID	NM_016955.4
cDNA	c.1321G>A
Protein effect	p.G441R
GnomAD allele frequency	Not present.
ClinVar status	No data available.
ACMG-Classification	Likely pathogenic: 7 points = 7 P – 0 B
PhastCons100way	1.000
PhyloP100way	7.108
Index II.1	T/T
Brother II.2 (affected)	T/T
Mother I.2 (healthy)	C/T
Father I.1 (healthy)	C/T

In family 81 we identified a homozygous missense variant in the *SEPSECS* gene at the genomic position chr4:25125738 that segregates within the family and is further described in Table 25. Mother and father are heterozygous carriers of this variant, the affected children are both homozygous carriers. The family pedigree is shown in Figure 15. This missense variant leads to a switch from cytosine to thymine and subsequently the amino acid glycine is changed to arginine (Gly441Arg).

It is classified as likely pathogenic and is located in a highly conserved area (PhastCons100way 1.000, PhyloP100way 7.108). It is not present in gnomAD.

This is a consanguineous family with the parents being first degree cousins. There are two affected children. The index patient II.1 showed an early-onset mental and intellectual development retardation. Her motoric skills worsened when she was in her early twenties. She showed balance impairment as well as gait difficulties. Her physical appearance is much younger and smaller than expected. At the age of 23 she developed rapid progressive atactic symptoms. The cMRI shows a significant cerebellar atrophy, but no thin corpus callosum and no signaling abnormalities as in ARSACS or leukoencephalopathy. According to medical history, her brother II.2 suffers from a mental retardation as well but he shows no physical impairment.

3.2.6 *NGLY1* – Family 88

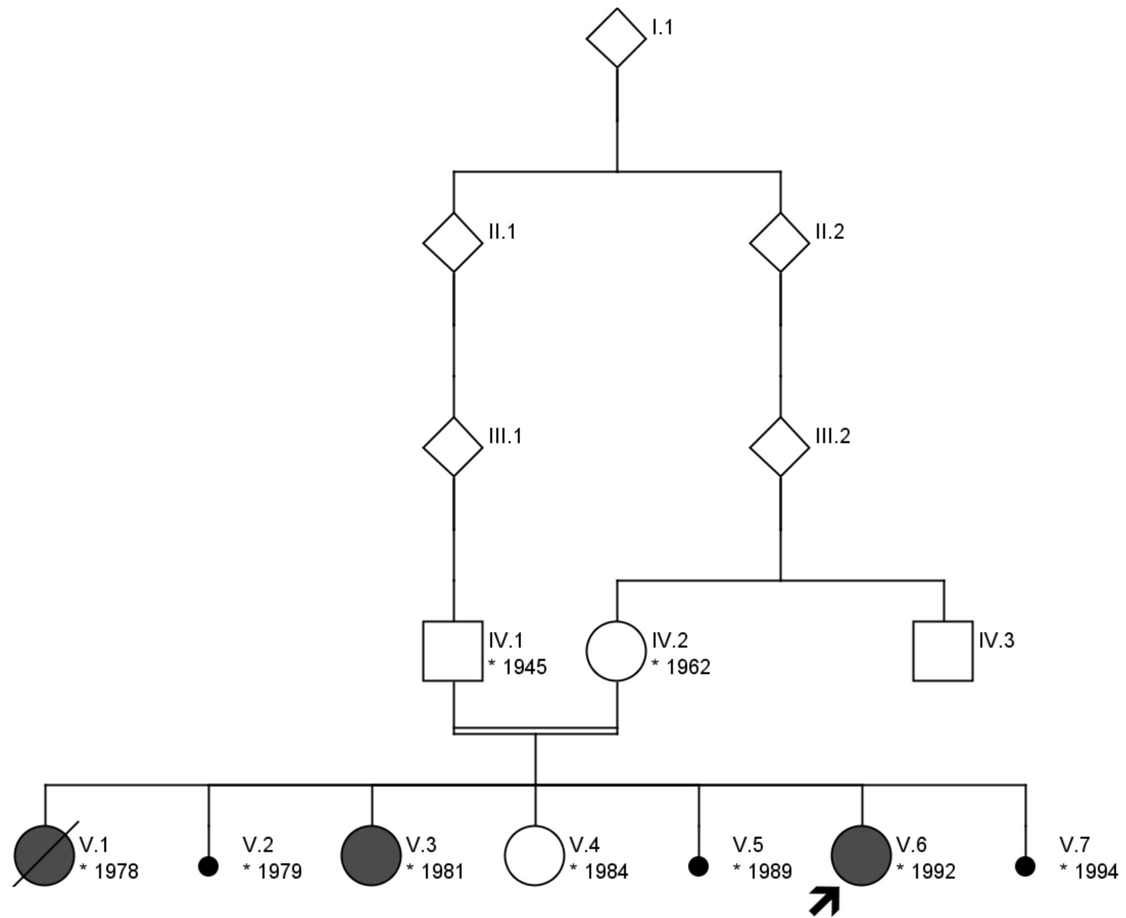

Figure 16: Pedigree of family 88.

Table 26: Overview of the NGLY1 variant.

Chromosome	3
Position	25770698
Variant type	INDEL
Reference	TTTGA
Alternate Allele	T
Zygosity	homozygous
Gene	*NGLY1*
Genbank transcript ID	NM_018297.4
cDNA	c.1533_1536delTCAA
Protein effect	p.N511Kfs*51

GnomAD allele frequency	Not present.
ClinVar status	Pathogenic
ACMG-Classification	Pathogenic: 17 points = 17 P – 0 B
Index IV.6	deletion/deletion
Mother III.2 (healthy)	wildtype/deletion
Uncle III.3 (healthy)	wildtype/deletion

In family 88 a truncating homozygous variant was found in the *NGLY1* gene. At the chromosomic position Chr3:25770698 the bases TTTGA were deleted, as described in Table 26. This deletion leads to a preterminal stop codon with the protein being shortened by 94 amino acids. It is classified as being pathogenic by the ACMG classification and is not present in gnomAD. Its pathogenic potential is supported by the fact that it is a homozygous truncating variant.

Its segregation within the family was confirmed by PCR and subsequent Sanger sequencing. The index patient carries a homozygous deletion, her healthy mother and uncle carry the wildtype on one allele and the deletion on the other allele.

As shown in Figure 16 this is a consanguineous family with the mother and the father of the index patient being second degree cousins. There are four daughters. Besides the index patient there are two more daughters who are also affected showing the same phenotype. Only one daughter appears unaffected and healthy. Furthermore, the mother had 4 abortions.

3.2.7 *MTPAP* – Family 16

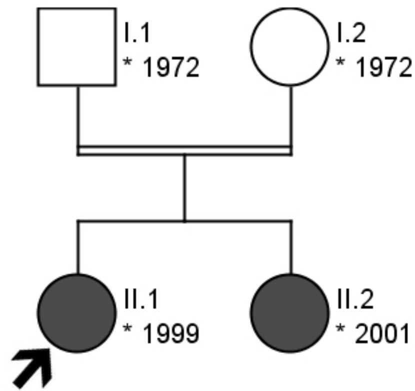

Figure 17: Pedigree of family 16.

Table 27: Overview of the MTPAP variant.

Chromosome	10
Position	30629278
Variant type	SNV
Reference	G
Alternate Allele	A
Zygosity	homozygous
Gene	*MTPAP*
Genbank transcript ID	NM_018109.4
cDNA	c.432C>T
Protein effect	p.Phe144=
GnomAD allele frequency	Variant not found.
ClinVar status	No data available.
ACMG-Classification	Likely benign: -1 points = 1 P – 2 B
PhastCons100way	1.000
PhyloP100way	1.832
Index II.1	A/A
Sister II.2 (affected)	A/A
Mother I.2 (healthy)	G/A
Father I.1 (healthy)	G/A

In family 16 we identified a homozygous missense variant in the *MTPAP* gene that segregates within the family. As described in Table 27 there is a switch from the base guanine to adenine at the genomic position chr10:30629278 that does not result in a change of the amino acid sequence. The ACMG-classification indicates that this variant is likely benign. It is well conserved (PhastCons100way 1.000, PhyloP100way 1.832).

As shown in Figure 17 this is a consanguineous Turkish Kurdish family with the parents being first degree cousins. They have two daughters who both suffer from an early onset ataxia. The age of onset in the index patient is 4 years, her sister showed first signs at the age of 3.

In order to further examine the variant we found in family 16 and to evaluate its pathogenic potential, we also performed mitochondrial poly (A) tail assays. Even though the amino acid sequence is not changed by this base exchange, it is possible that the stability of the amino acid chain is impaired and therefore the polyadenylation is affected negatively.

For the polyadenylation assay we examined the mRNA of our index patient II.1, her affected sister II.2 and a control mRNA that appears healthy and is of similar age as our patients. Furthermore, we also treated two batches of our control fibroblasts with Cordycepin in order to inhibit the polyadenylation. (Wong *et al.*, 2010)

The fragment analysis of our probes did not show any difference in the polyadenylation status (Figure 18). Neither in the affected patients nor in the treated control probes.

Therefore, we assume that our identified variant is not disease causing and family 16 remains genetically unsolved.

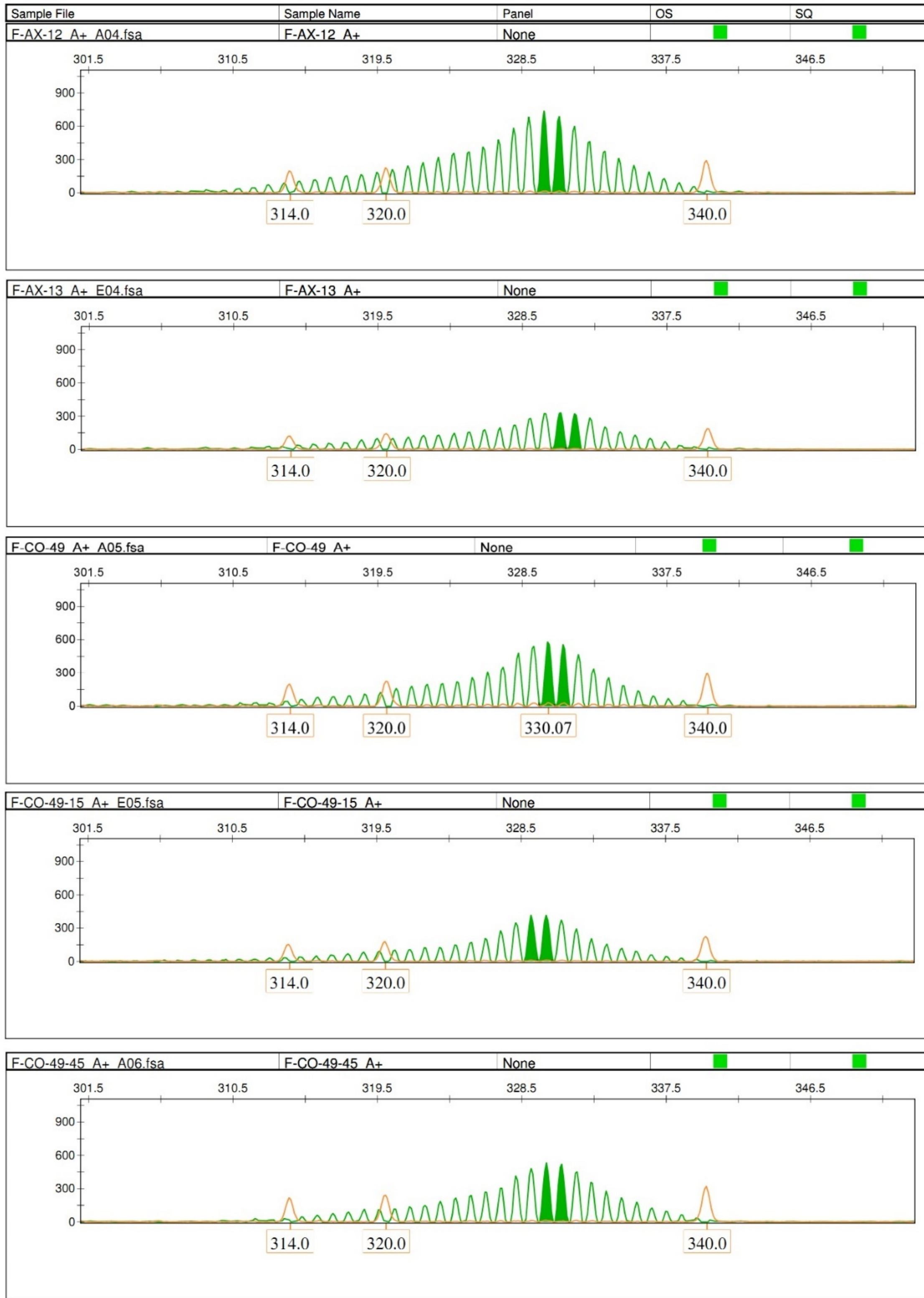

Figure 18: Fragment analysis of the polyadenylation assay. There is no difference in the polyadenylation status between the affected patients and the healthy control patient.

3.3 Solved by other laboratories

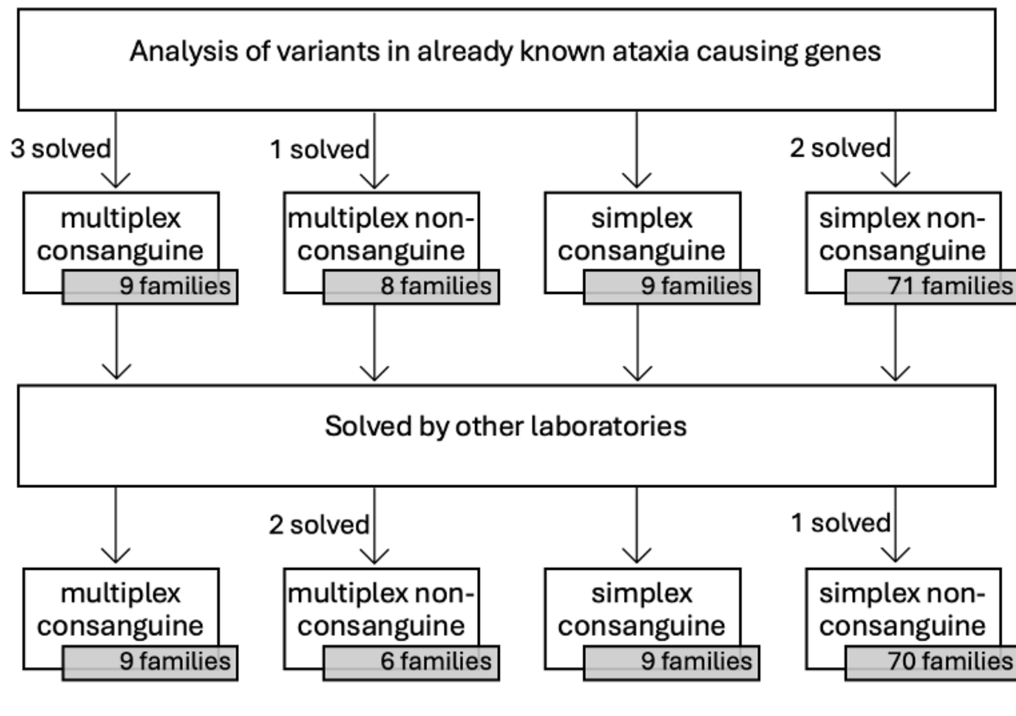

Figure 19: After analysing all families for variants in already known ataxia causing genes, 3 more families were solved by other laboratories.

As shown in Figure 19 three families were genetically diagnosed by other laboratories.

Table 28: Overview of the variants in known ataxia genes that were found by other laboratories.

Family	Chromosome	Position	Reference	Alternate	Genotype	Gene	Variant type
26	1	1451415	C	G	C/G	*ATAD3A*	Missense
26	1	1452793	G	A	G/A	*ATAD3A*	Splice acceptor variant
39	MT	8881	T	C	T/C	*MT_ATP6*	Missense
69	9				>1000 GAA Repeats	*FXN*	Repeat Expansion

Table 28 summarises the variants that were found by other laboratories in genes that are already known to cause ataxia.

3.3.1 *ATAD3A* – Family 26

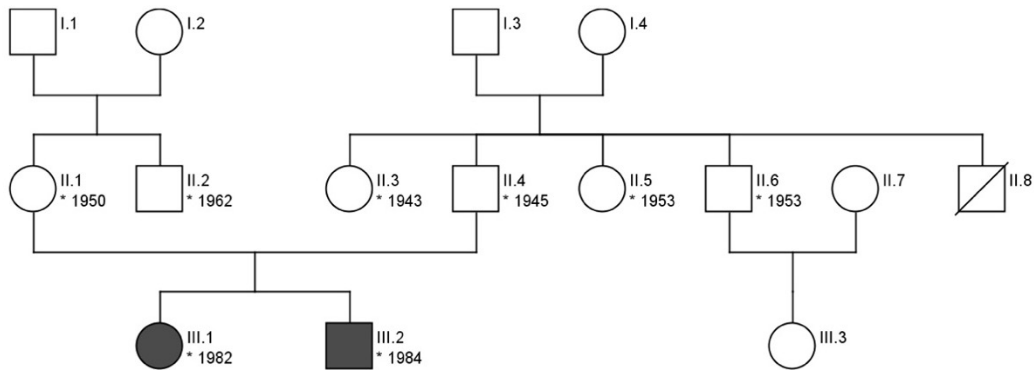

Figure 20: Pedigree of family 26.

Table 29: Overview of the ATAD3A variant.

Chromosome	1	1
Position	1451415	1452793
Variant type	Missense	Splice acceptor variant
Reference	C	G
Genotype	G	A
Gene	*ATAD3A*	*ATAD3A*
Genbank transcript ID	NM_001170535.3	NM_018188.5
cDNA	c.229C>G	c.528+1G>A
Protein effect	p.Leu77Val	g.5263G>T
GnomAD allele frequency	Not present.	Not present.
ClinVar status	Uncertain significance.	Likely pathogenic.
ACMG-Classification	Uncertain significance: 4 points = 5 P – 1 B	Likely pathogenic: 9 points = 9 P – 0 B
PhastCons100way	0.968	1.000
PhyloP100way	2.021	9.266
Sister III.1 (affected)	C/G	G/A
Brother III.2 (affected)	C/G	G/A

Concurrently with this study, the Institute of Medical Genetics and Applied Genomics at the University Hospital Tübingen has detected two variants in the *ATAD3A* gene, described in Table 29, which segregate in family 26.

First there is a heterozygous missense variant at the genomic position chr1:1451415 with a switch from cytosine to guanine leading to a change in the amino acid chain from leucine to valine. The clinical variant status is of an unknown significance. According to PhastCons100way and PyhloP100way this position is well conserved.

The second variant is a heterozygous single base exchange in the intronic region. It does not lead to a change on protein level.

In this non consanguineous family, as shown in Figure 20, there are two affected siblings with an early onset neurological disorder. At the initial presentation the male patient showed leg spasticity, external ophthalmoplegia, early childhood cataract, extrapyramidal symptoms and severe mental retardation. The accompanying parents reported that the first surgery was performed at the age of 5 months due to a bilateral cataract. They furthermore reported delayed cognitive and motor development. He learned to walk and speak his first two-word sentences at 2 years old. At the age of 3 years, he was diagnosed with an intellectual disability. He learned to climb stairs at about the age of 5 years. His fine motor skills are also very limited since childhood.

His older sister presented with cerebellar ataxia with mild cognitive impairment, early childhood cataract and external ophthalmoplegia. She first showed symptoms at the age of 6 years. She has a slowly progressive gait and fine motor and a mild cognitive impairment. Her parents report that she learned to walk independently at the age of 14-18 months. The development of speech had started early, she talked and sang a lot and was also active in motor activities, for example, she climbed a lot. At the age of 6 years, she walked noticeably unsteadily and over time, the gait pattern and coordination had further deteriorated. Furthermore, cognitive impairments as well as psychiatric abnormalities had been noticed early on with short attention span, depression and impulse and aggression control disorder. Further follow-up examinations showed the disease to be progressive, as she was requiring walking aids and developing hypacusis. A cMRI was performed in 2009 and showed a vermal accentuated cerebellar atrophy.

3.3.2 *MT_ATP6* – Family 39

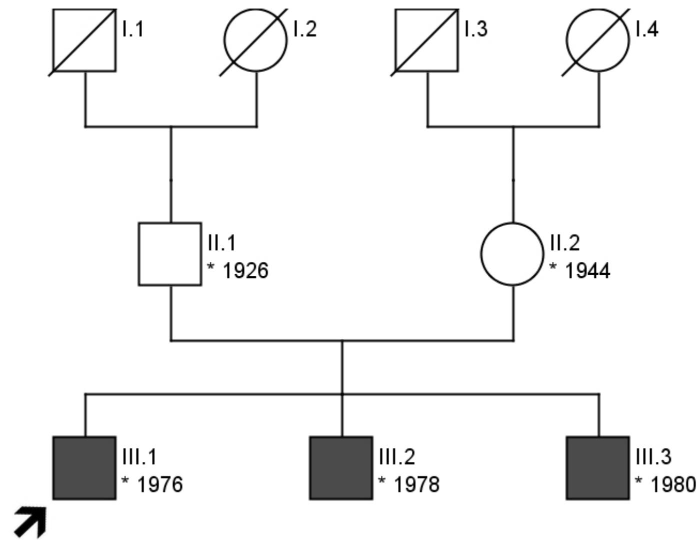

Figure 21: Pedigree of family 39.

Table 30: Overview of the MT_ATP6 variant.

Chromosome	MT
Position	8881
Variant type	SNV
Reference	T
Genotype	C
Gene	*MT_ATP6*
Genbank transcript ID	NC_012920.1
cDNA	c.355T>C
Protein effect	p.Ser119Pro
gnomAD mitochondrial	Variant not found.
MitoMap	0; likely pathogenic
ClinVar status	No data available.
ACMG-Classification	Uncertain significance: 5 points = 5 P – 0 B
PhastCons100way	0.000
PhyloP100way	0.576

In family 39 molecular genetic testing revealed the presence of a variant in the *MT-ATP6* gene at the position m.8881T>C leading to a p.Ser119Pro switch as shown in Table 30. We were not able to detect this variant using whole genome sequencing or whole exome sequencing because this gene is part of the mitochondrial genome and the Genesis analysis pipeline was not suited to detect mitochondrial variants.

In this non consanguineous family, there are three affected male siblings with a neurological disorder as shown in Figure 21. The index patient initially presented himself at the neurological department at the age of 35. He reported of an early onset cerebellar ataxia with sensorimotor axonal polyneuropathy, pyramidal tract lesion, discrete cognitive performance impairment, action-depending myoclonia and bladder and fecal incontinence. In 2009 a cMRI was performed which shows cerebellar atrophy of a cerebellum that is probably already primarily hypoplastic without a thin corpus callosum or myelon atrophy. His two younger brothers also suffered from ataxia. His parents and other family members appeared to be healthy.

3.3.3 *FXN* – Family 69

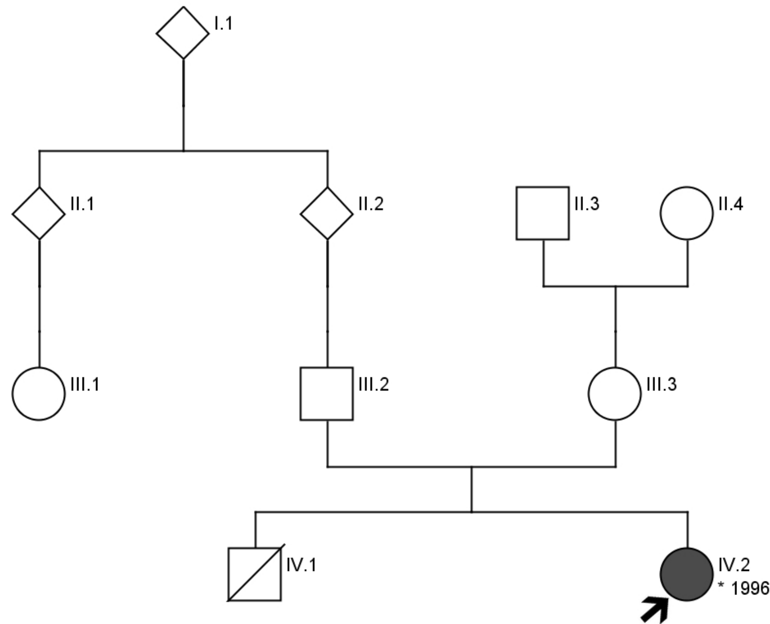

Figure 22: Pedigree of family 69.

In family 69 molecular genetic testing was performed and a disease-causing homoallelic GAA expansion with about 1000 repeats in the *FXN* gene was verified.

As shown in Figure 22 this is a non-consanguineous family. The index patient presented with cerebellar ataxia since the age of 12 years. Symptoms began insidiously with uncoordinated gait and balance but without involvement of fine motor skills. On initial neurologic examination, sluggish speech was noted. She also showed a positive Babinski sign on both sides and hollow feet. Regular follow-up examinations showed an increase in sluggish speech but no reduction in walking distance.

3.4 New candidate genes

After screening all families for variants in previously known ataxia-causing genes, our aim was to identify new disease-causing genes. Therefore, the patients' genomes were analysed again with Genesis. Filter criteria now included a minor allele frequency of 0.2% (MAF in ExAC, 1000 Genomes and EVS <0.2%).

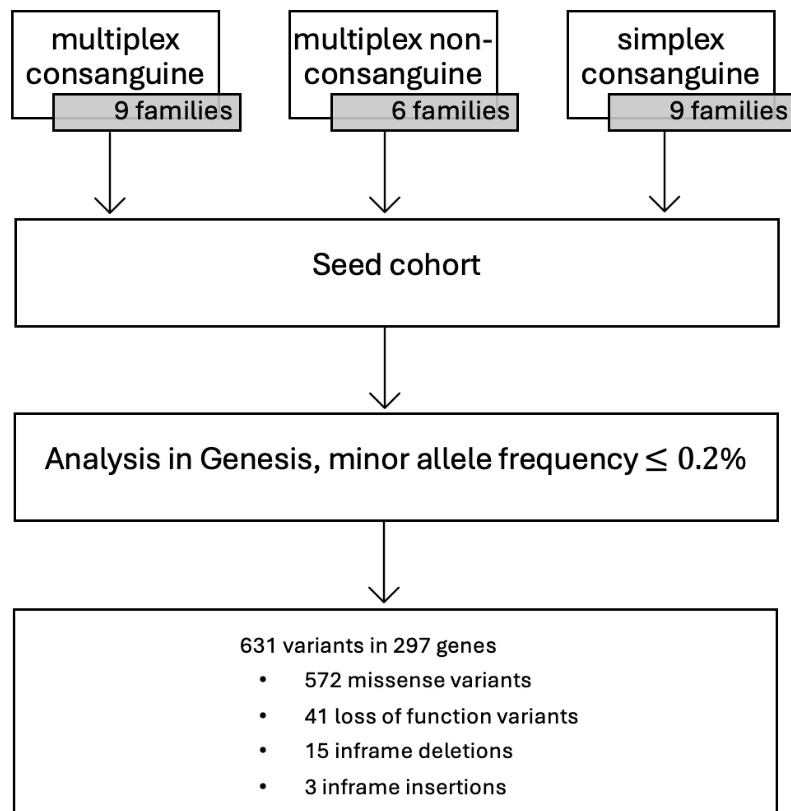

Figure 23: Forming and analysing the seed cohort consisting of all multiplex and simplex consanguine cases using Genesis.

As shown in Figure 23 first the seed cohort was analysed in Genesis. 631 variants (572 missense variants, 41 loss of function variants, 15 inframe deletions, 3 inframe insertions) in 297 genes were found.

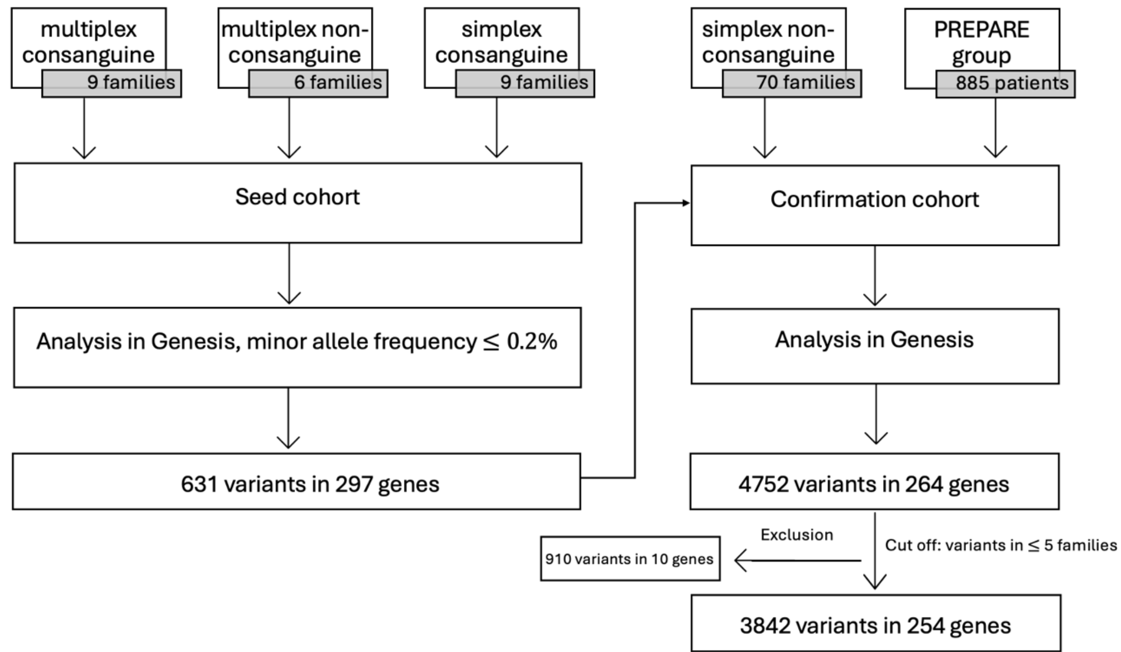

Figure 24: The search for new genes started with an analysis in the seed cohort in Genesis. The identified variants were then rechecked in the confirmation cohort.

As shown in Figure 24 we rechecked those newfound genes in the confirmation cohort to find out if there are other families also having variants in one of those genes to increase the likelihood of identifying new disease-causing variants. 4752 variants (4227 missense variants, 351 loss of function variants, 2 variants in coding sequence, 106 inframe deletions, 61 inframe insertions, 4 protein altering variants and 1 stop retaining variant) in 264 genes were found. Having a closer look at these hits a cutoff was set to only include variants in the final table that occurred in a maximum of 5 families. With that 910 variants in 10 genes were eliminated from the final table.

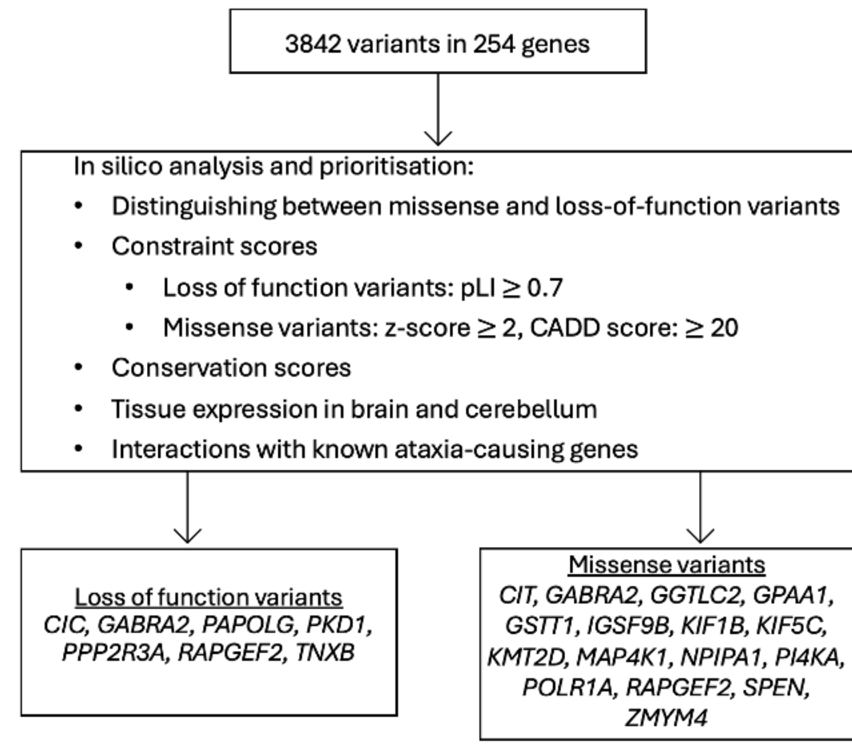

Figure 25: Further in silico analysis and prioritisation led to the identification of 21 potential new candidate genes.

The resulting genes were analysed regarding their pathogenic potential, evolutionary conservation, tissue distribution and possible interaction with already known ataxia-causing genes as shown in Figure 25.

Using this approach, we identified 21 new potential ataxia-causing genes in silico.

Table 31: Overview of the 21 new potential ataxia-causing genes identified in silico in the seed cohort.

Family	Chromosome	Position	Reference	Alternate	Genotype	Gene	Variant type
92	19	42773535	T	G	T/G	CIC	Splice donor
92	19	42776500	C	A	C/A	CIC	Missense
92	19	42799164	A	C	A/C	CIC	Missense
63	12	120270646	G	C	C/C	CIT	Missense
92	4	46388157	TG	T	TG/T	GABRA2	Frameshift
92	4	46388160	T	A	T/A	GABRA2	Missense
47	22	22989241	G	C	G/C	GGTLC2	Missense

47	22	22989298	A	G	A/G	GGTLC2	Missense
43	8	145139449	C	T	T/T	GPAA1	Missense
62	8	145138868	G	A	A/A	GPAA1	Missense
18	22	24376825	C	T	T/T	GSTT1	Missense
37	11	133789656	G	C	C/C	IGSF9B	Missense
71	1	10322003	A	G	A/G	KIF1B	Missense
71	1	10338081	G	T	G/T	KIF1B	Missense
16	2	149866716	C	T	T/T	KIF5C	Missense
16	12	49420856	C	T	C/T	KMT2D	Missense
16	12	49432470	G	C	G/C	KMT2D	Missense
47	19	39079854	A	G	A/G	MAP4K1	Missense
47	19	39079859	C	A	C/A	MAP4K1	Missense
37	16	15045754	C	G	C/G	NPIPA1	Missense
37	16	15031321	G	A	G/A	NPIPA1	Missense
37	16	15045859	C	T	C/T	NPIPA1	Missense
20	2	60987307	A	G	G/G	PAPOLG	Missense
20	22	21083710	C	G	G/G	PI4KA	Missense
16	16	2156401	G	A	A/A	PKD1	Missense
32	2	86280095	C	T	T/T	POLR1A	Missense
70	3	135789340	TG	T	TG/T	PPP2R3A	Frameshift
70	3	135789342	G	C	G/C	PPP2R3A	Missense
107	4	160262952	T	C	T/C	RAPGEF2	Missense
107	4	160262959	C	G	C/G	RAPGEF2	Missense
71	1	16259904	C	T	C/T	SPEN	Missense
71	1	16262073	G	A	G/A	SPEN	Missense
37	6	32049177	C	T	C/T	TNXB	Missense
37	6	32065101	G	A	G/A	TNXB	Missense
52	1	35824553	C	G	G/G	ZMYM4	Missense

Following a thorough literature review of these genes and their variants, shown in Table 31, our focus turned to the *GPAA1* gene. The other in Table 29 listed variants were not analysed.

The *GPAA1* gene is significantly involved in GPI anchor synthesis and other GPI-associated diseases associated with intellectual disability, hypotonia and epilepsy have already been described in the literature. By that time, it was not yet described as being an ataxia-causing gene. In two families of our seed cohort homozygous missense variants in the *GPAA1* gene were identified.

We found four patients in two consanguineous families carrying homozygous variants in the *GPAA1* gene. In both families the patients presented with an early onset cerebellar ataxia with quite similar phenotypes including cerebellar atrophy, mental retardation, developmental delay and the inability to walk independently.

3.4.1 *GPAA1* – Family 43 and 62

<u>Family 43</u>

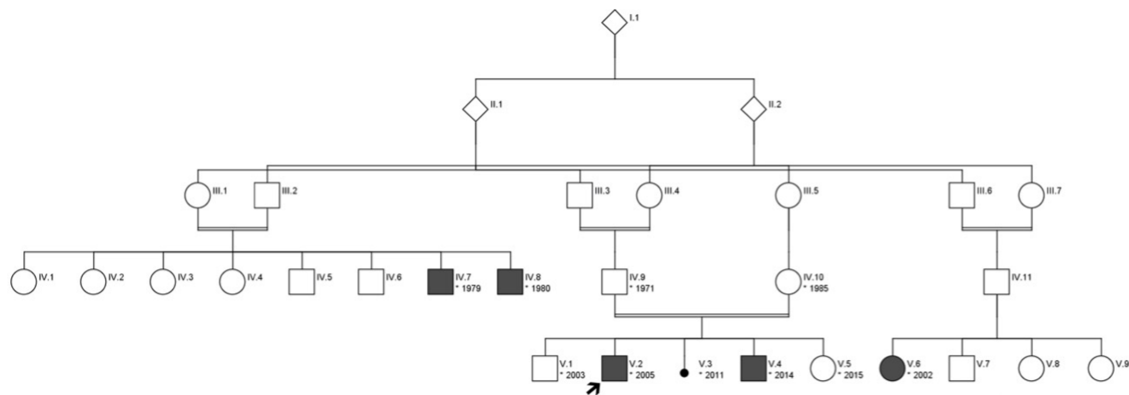

Figure 26: Pedigree of family 43.

As shown in Figure 26 family 43 is a consanguineous family with 5 affected persons. The index patient had experienced significant developmental delays from birth but had never regressed. There was a history of seizures in infancy, although these could not be described in detail. He started to crawl at 24 months of age. He could sit unaided from the age of 36 months. He can only walk with the aid of a rollator; otherwise, his gait is

very wobbly and ataxic. His first words were spoken at the age of four, and his vocabulary is not age appropriate. His speech also appears dysarthric. He has an alternating convergent strabismus. There are saccadic eye movements and in general a reduced muscle tone. Dysmetria and an intention tremor are also present. The cMRI shows clear cerebellar atrophy, which is progressive as shown by follow-up examinations.

<u>Family 62</u>

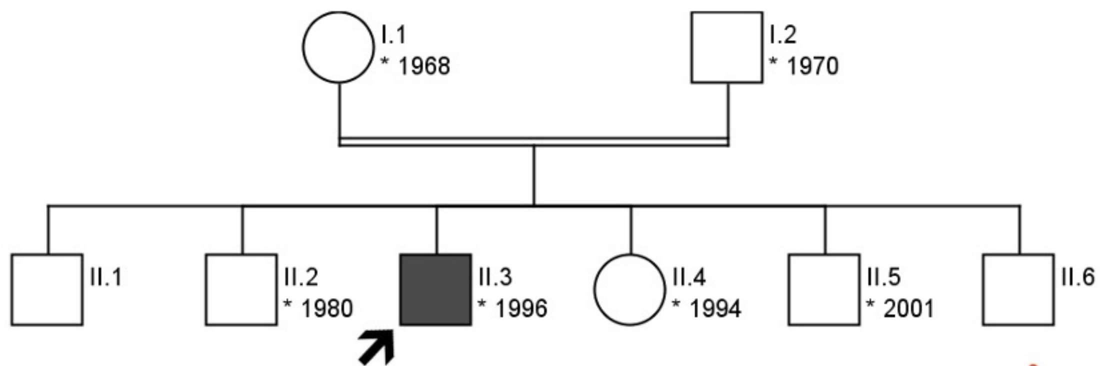

Figure 27: Pedigree of family 62.

Family 62 is also consanguineous with one affected person and 5 healthy siblings as shown in Figure 27. The index patient has a developmental delay, coordination disorder and learning disability. The gait pattern is clumsy and the coordination of both the upper and lower extremities appears limited. There is also moderate diadochokinesis. The muscle tone appears hypotonic. The cMRI shows pronounced cerebellar atrophy. His facial features appear dysmorphic, characterised by bulging lips, a long face, and myopathic facies.

Whole exome sequencing was performed in both family 43 and family 62 and no variant in the known ataxia-causing genes could be detected.

We identified two homozygous missense variants in the *GPAA1* gene.

Table 32: Overview of the GPAA1 variants in family 43 and 62.

Family	43	62
Chromosome	8	8
Position	145139449	145138868
Variant type	SNV	SNV
Reference	C	G
Alternate Allele	T	A
Zygosity	homozygous	homozygous
Gene	*GPAA1*	*GPAA1*
Protein Notation	ENSP00000347206.4: p.Ala316Val	ENSP00000347206.4: p.Val181Ile
Genbank transcript ID	NM_003801.4	NM_003801.4
cDNA	c.947C>T	c.541G>A
Protein effect	p.A316V	p.V181I
GnomAD allele frequency	Not present.	Not present.
ClinVar status	Uncertain significance.	Likely benign
ACMG-Classification	Uncertain significance: 5 points = 5 P – 0 B	Likely benign: -5 points = 1 P – 6 B
PhastCons100way	1.000	1.900
PhyloP100way	7.118	5.540

As shown in Table 32 the first variant is located at the genomic position chr8:145139449 with a change from cytosine to thymidine resulting in an alteration in the amino acid chain from alanine to valine. This site is highly conserved (PhastCons100way 1.000, PhyloP100way 7.118). It is classified with an uncertain significance. It is not present in gnomAD.

The second variant is at the genomic position chr8:145138868. It is a single-nucleotide variant with guanine being exchanged for alanine resulting in a change in the amino acid chain. Isoleucine is now inserted in place of valine. The conservation scores are a little

bit lower than for the first variant, but nevertheless it shows a high conservation (PhastConsway100 1.900, PhyloP100way 5.540). This variant is classified as likely benign and it also is not present in gnomAD.

Table 33: Genetic analysis in family 43 and family 62..

Family 43		Family 62	
Index V.2	T/T	Index II.3	A/A
Mother IV.10 (healthy)	C/T	Mother I.2 (healthy)	G/A
Father IV.9 (healthy)	C/T	Father I.1 (healthy)	G/A
Brother V.1 (healthy)	C/T	Brother II.2 (healthy)	G/G
Brother V.4 (affected)	T/T	Sister II.4 (healthy)	G/A
Uncle IV.11 (healthy)	C/T	Brother II.5 (healthy)	G/A
Cousin V.6 (affected)	T/T		

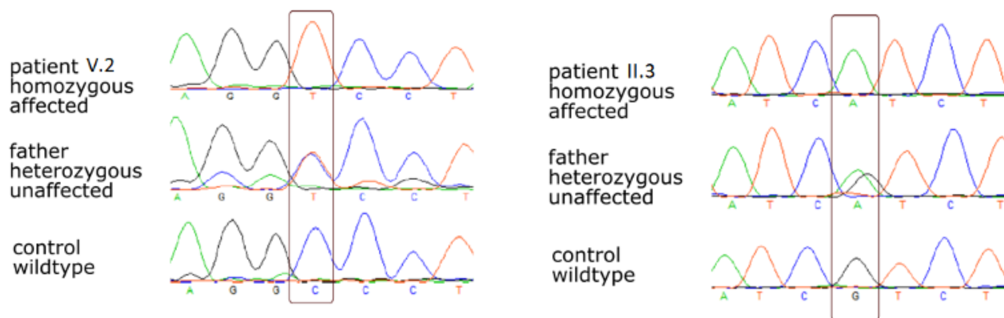

Figure 28: Sanger sequencing analysis of the affected homozygous index patient, one unaffected heterozygous family member and an unaffected wildtype control for each of the two families.

As shown in Table 33 and Figure 28 both variants were validated via PCR and subsequent Sanger sequencing and are segregating within the respective family in an autosomal recessive manner.

The *GPAA1* gene encodes for an anchor attachment protein. If our variants decrease the number of functioning GPI anchors less proteins will be bound to the cell surface. Therefore, to further validate the pathogenicity of the above shown variants we quantified

the abundance of GPI anchored proteins on the surface of patients' peripheral blood monocytes compared to controls.

Quantification was performed by our college Prof. Möhle, Department of Hematology and Oncology, Medical Center, University Hospital of Tübingen, Germany, using the FLAER based flow cytometry. FLAER is the abbreviation of fluorescein-labeled proaerolysin. This flow cytometry analysis uses antibodies that are directed against GPI anchored proteins. It was performed with peripheral blood monocytes of our index patient V.2 and a healthy control.

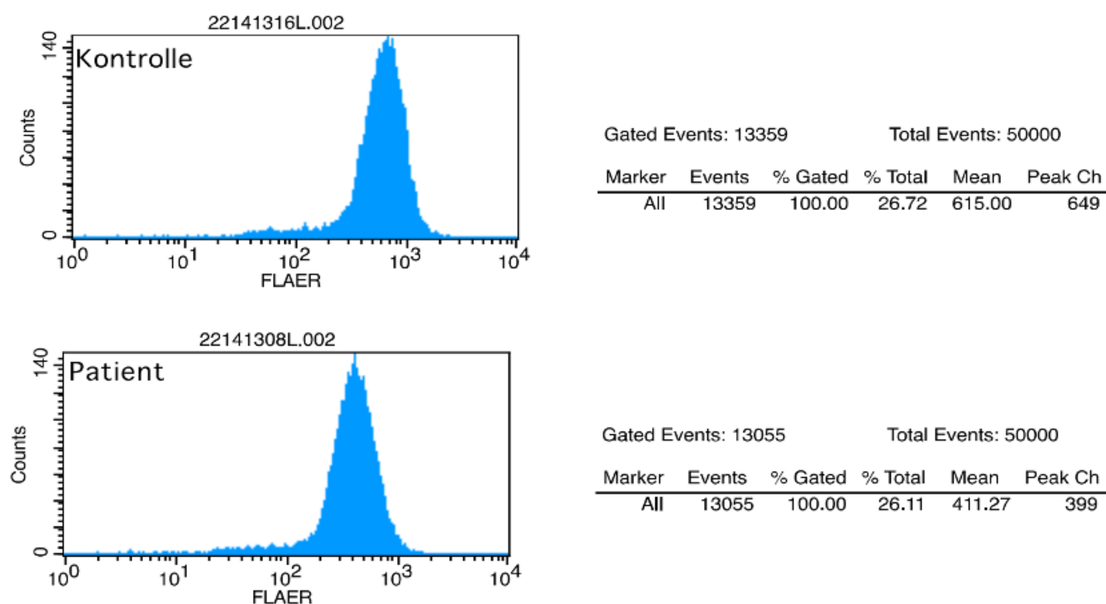

Figure 29: The surface abundance of FLAER, a marker for GPI anchor proteins, was measured on control (upper image) and patient (lower image) PBMCs by FACS analysis. Patient fibroblasts show a reduction of GPI anchor protein expression on the cell surface by ~30%.

The results show a significant reduction of GPI anchored proteins on the cell surface by approximately 30% as shown in Figure 29. This leads us to the conclusion that variants in the *GPAA1* gene can cause a reduced expression of GPI anchored proteins on the cell surface leading to an early onset cerebellar ataxia.

4 Discussion

Hereditary ataxias appear with a variety of symptoms and are genetically very heterogenous. Despite the growing number of genetic testing methods available, molecular diagnosis remains challenging in many cases. In addition to known genes, rare or previously undescribed variants that further differentiate the genetic landscape of these diseases are increasingly appearing.

The aim of this study was to use modern sequencing technologies to identify new disease-causing variants in patients with cerebellar ataxias of unknown genetic origin, and to provide a genetic diagnosis for as many families with atactic symptoms as possible. For this purpose, both already known disease-causing genes and new candidate genes were investigated.

We focused particularly on recessive inheritance in consanguineous and multiplex families. Consanguineous families have a much higher risk of being affected by recessive inherited diseases because the likelihood that two disease causing recessive variants are being inherited is two times higher compared to non-consanguineous families. In the general population there is a 2-4% likelihood to inherite an unpredicted disease. In consanguineous families where first degree cousins are married the likelihood doubles. (Murken, 2006)

Autosomal recessive genetic diseases can be caused by both homozygous and compound heterozygous variants. Homozygosity for a variant occurs when both alleles of a gene carry the same variant. If both copies of a gene contain the same disease-causing variant, the gene loses all function and the disease manifests itself. Compound heterozygosity is present if two different disease-causing variants are present in the same gene, on each of the two alleles. Although these variants are distinct, they both result in a loss or severe restriction of gene function. Since there is no normal copy, the gene product cannot be produced in sufficient quantity or quality, resulting in the development of the disease. Both homozygous and compound heterozygous variants can lead to disease if both alleles severely impair the function of a gene to the extent that its normal physiological role can no longer be fulfilled.

A variant is a permanent change to an organism's DNA sequence that can affect protein biosynthesis, influence the function of the encoded protein and change the resulting phenotype. The structure of a protein is closely linked to its function. Proteins consist of a specific sequence of amino acids that fold into a particular three-dimensional shape. Even a single amino acid change, such as a point variants, can influence this folding, particularly if the new amino acid has different chemical properties, such as charge, polarity or size. If the amino acid change affects the active site of an enzyme, the enzyme may no longer recognise its substrate or may be unable to convert it effectively. Similarly, if a variant leads to an amino acid change in the binding site of a protein, its binding capacity may be reduced or eliminated entirely. Furthermore, an unsuitable amino acid can cause the protein to become unstable, resulting in it being broken down more quickly or forming aggregates in the body.

Of the 103 families with hereditary autosomal recessive cerebellar ataxia in the studied cohort, disease-causing variants were found in 6 families in the known ataxia-associated genes *AFG3L2, NGLY1, PTRH2, RNF216, SBF1* and *SEPSECS*. These genes are associated with a broad phenotypic spectrum that may include additional neurological symptoms alongside ataxia. These results emphasise the diagnostic value of exome and genome sequencing, even with regard to rare but known disease genes.

In 3 other families, variants in the *ATAD3A, FXN* and *MT-ATP6* genes were identified by collaborating laboratories. These findings provide a fuller picture of the genetic causes of hereditary cerebellar ataxias within the cohort and emphasise the importance of collaboration and data sharing between centres to optimise the diagnostic rate.

The fact that 9 families with variants in known ataxia genes could be solved may be due to the following factors:

1. The sensitivity of earlier tests was limited. Panel diagnostics only examine selected genes, and whole exome sequencing only covers the coding regions of the genome. This means that variants in regulatory regions or structural variants often remain undetected.
2. The technical limitations of past sequencing methods. Standard exome sequencing is not reliable for detecting repeat expansion disorders such as Friedreich's ataxia. Large deletions or duplications can also be overlooked. Older analyses may also

fail to detect variants with a low allele frequency that fall below the detection limit. The potential of today's genome sequencing lies in its ability to detect non-coding regions, structural variants and mitochondrial DNA. This capability surpasses that of conventional panel diagnostics or exome sequencing, enabling the identification of variants that may have been overlooked.

3. Advances in bioinformatics and interpretation. Older analyses may have classified variants as 'variants of unclear significance', but more recent findings have since reclassified them as pathogenic. Better interpretation can also be achieved through changes in annotation, databases and algorithms. Therefore, reanalysing older sequence data can now lead to a diagnosis.
4. The variant had previously been overlooked or misclassified. It is possible that the variant was present in the datasets but went unrecognised due to incorrect filtering, poor exome coverage or an unclear genotype-phenotype association at the time.

Two missense variants in the *GPAA1* gene were identified in two unrelated families that were undescribed at the time of data collection in this study.

GPAA1 is a gene that encodes a subunit of the GPI transamidase complex. This complex is essential for the post-translational attachment of GPI anchors to proteins, which is a central process to the anchoring of proteins to the cell membrane.

Segregation analyses demonstrated autosomal recessive inheritance. Functional validation was performed using FLAER analysis, which revealed decreased expression of GPI-anchored proteins on the cell surface. This supports the functional relevance of the variants and strongly indicates the pathogenicity of the variant.

Previously published data suggest possible involvement in syndromic diseases. The cases described here are not isolated ataxia either, since in addition to cerebellar symptoms, the patients also have impaired cognitive abilities and a developmental disorder.

Our findings suggest that *GPAA1* is another gene that should be considered in cases of syndromic cerebellar ataxia, particularly when it is accompanied by cognitive impairment. Further research involving larger patient cohorts and functional analyses is required to determine the precise function of *GPAA1* in neurodegenerative diseases.

If only diagnoses determined by our sequencing are taken into account, the genetic clarification rate achieved here is 7.8% (eight out of 103 families). Including the externally resolved cases brings the overall clearance rate to 10.7% (11 out of 103 families). This diagnosis rate appears to be lower than that achieved in similar studies. Depending on factors such as the sequencing method used, selection criteria, clinical phenotyping and the ethnic composition of the cohorts, diagnosis rates in the literature vary between 20% and 50%. (Anheim *et al.*, 2010; Németh *et al.*, 2013; Fogel *et al.*, 2014; Ngo *et al.*, 2020)

The comparatively low rate of our study could be due to several factors:

1. The heterogeneity of phenotypic presentation, particularly in cases involving additional neurological or cognitive symptoms, often makes targeted gene prioritisation challenging.
2. A high proportion of previously unknown or difficult to interpret variants in new or less characterized genes. Even with high technical effort, the interpretation of variants of unknown significance remains limited if functional validation is not possible.
3. Possible selection bias: many patients in the cohort had already undergone previous genetic testing, so this could be a residual cohort that is more difficult to diagnose.
4. Technical limitations such as low detection of structural variants or non-coding variants despite genome sequencing could also play a role.

The above factors provide a good starting point for improving the diagnosis rate. Using deep phenotyping for phenotypic-genetic feedback to enable more precise gene filtering through standardised symptom input, or reverse phenotyping to target retesting of the candidate gene and determine whether the phenotype matches, would be beneficial.

Furthermore, the diagnosis rate can be increased by re-analysing data using current databases, as variants that were previously unclear may now be classified as disease-causing. Therefore, functional validation is necessary to prove the pathogenicity of variants.

Improvement in sequencing technology, such as the us of whole genome sequencing instead of whole exome sequencing, long-read sequencing, which better covers complex genetic regions, repeats and structural variants, and deep sequencing, which increases sensitivity, would also support a higher diagnosis rate.

This study highlights the genetic heterogeneity of hereditary cerebellar ataxias, demonstrating the importance of comprehensive sequencing technology combined with functional analysis. Identifying pathogenic variants in *GPAA1* expands the range of known autosomal recessive cerebellar ataxia genes, suggesting that defects in GPI biosynthesis may play a more significant role than previously thought in syndromic forms of cerebellar ataxia.

The study also has some limitations. Firstly, the small number of *GPAA1*-positive families (two cases) makes it difficult to draw conclusions about frequency and phenotypic variability. Secondly, functional characterisation was performed exclusively by FLAER analysis. While this provides evidence of reduced GPI-anchored protein expression, further studies, such as the use of animal models, are needed to improve our understanding of the disease-relevant mechanisms. Furthermore, potential environmental factors or genetic modifiers that could influence clinical expression have not been systematically investigated.

It is very well possible that some families could not be solved, because here no recessive inheritance is causative, but for example an X-linked or a mitochondrial inheritance. In non-consanguineous simplex families, a de novo variant is also quite possible.

Overall, this work expands the genetic spectrum of hereditary autosomal recessive cerebellar ataxias by confirming new and rare genes that are already known to cause the disease. The results presented here emphasise the importance of personalised diagnostic strategies and could enable targeted therapy approaches in the long term.

Future studies should include in-depth phenotype-genotype correlations, systematic functional analyses, and international collaborations to identify additional rare genes, with the aim of improving the diagnostic rate and identifying therapeutic targets.

4.1 Known genes

4.1.1 *AFG3L2* – Family 29

In family 29 we identified two heterozygous missense variants in the *AFG3L2* gene that segregate within the family.

The *AFG3L2* gene encodes for the ATPase Family Gene 3 Like Matrix AAA Peptidase Subunit 2. This is an enzyme that is located at the inner mitochondrial membrane, interacts with proteins and the ribosome and is very closely related to paraplegin. Variants in this gene are already known to cause dominant spinocerebellar ataxia 28, recessive spastic ataxia 5 and dominant optic atrophy. (Koppen *et al.*, 2007; Baderna *et al.*, 2020; Caporali *et al.*, 2020)

In 2020 functional analyses were performed showing that the two found heterozygous missense variants trigger optic atrophy and a spastic ataxia. Carriers of only the p.A462V variant are the mother and brother of our index patient. Showing a pure optic atrophy, they fit the published phenotype.

The index patient's father is a carrier of the p.Q620K variant. He also fits the phenotype and developed mild atactic symptoms later in life showing no signs of a visual impairment.

The index patient presented with a combined phenotype with optic atrophy and childhood onset progressive gait disturbance, spastic-ataxia and dystonia. She carries not only the p.A462V variant that is linked to the optic atrophy but also the p.Q620K variant that causes her atactic symptoms.

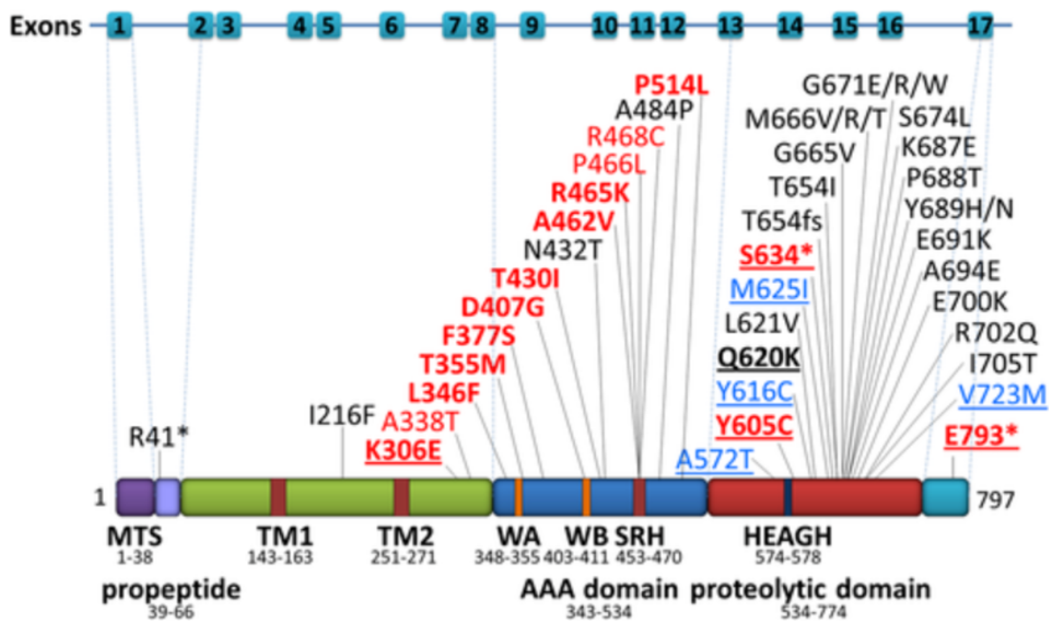

Figure 30: AFG3L2 protein with variant distribution and their phenotypical association (red = optic atrophy, black = spinocerebellar ataxia 28, blue = spastic ataxia syndrome 5). Source: Reproduced with permission from Annals of Neurology, John Wiley and Sons, "ATPase Domain AFG3L2 Mutations Alter OPA1 Processing and Cause Optic Neuropathy", Taroni et al., 2020, https://doi.org/10.1002/ana.25723

Interestingly the resulting phenotype is linked to the localisation of the variant. Optic atrophy seems to be associated with variants within the ATPase domain, whereas spastic and atactic symptoms occur with variants in the proteolytic domain (see Figure 30).

Using a yeast model, the variants pathogenicity was demonstrated. Further cell studies were performed, showing changes in *OPA1* that were also demonstrated visually due to fragmentation of the mitochondrial network within patients' fibroblasts. (Caporali *et al.*, 2020)

Current therapeutic research approaches include the use of Ca^{2+} modulators, such as ceftriaxone, to improve motor function. Administering the drug before the onset of symptoms prevented the occurrence of ataxia. (Maltecca *et al.*, 2015) Targeting proteotoxicity with chloramphenicol has also been shown to be effective in restoring the shape of mitochondria in *AFG3L2* variants in vitro. (Richter *et al.*, 2019) Further research into therapeutic approaches requires clinical studies on calcium modulators and, potentially, combination therapies. It would also be important to develop biomarkers that reflect mitochondrial stress in order to demonstrate the effectiveness of potential treatments at an early stage.

4.1.2 *PTRH2* – Family 53

In family 53 we identified a truncating and a missense variant in the *PTRH2* gene. Their segregation within the family was confirmed via PCR and subsequent Sanger sequencing.

The *PTRH2* gene encodes for the enzyme peptidyl-tRNA-hydrolase 2. It is a highly conserved gene that catalyses the hydrolysis of dissociated peptidyl-tRNA's which would accumulate within the cell and disturb or even inhibit the synthesis of new proteins. Furthermore, it is an important protein that regulates the cell cycle by being part of an integrin signaling complex and also mediating against apoptosis by inducing bcl-2. On the other hand PTRH2 also mediates towards anoikis by being phosphorylated when the cell is losing its attachment to the extracellular matrix. (Hu *et al.*, 2014)

The disease caused by the *PTRH2* variant is abbreviated IMNEPD. Written out this means infantile-onset multisystem neurologic, endocrine and pancreatic disease. Patients present with microcephalgia, mental retardation, hearing loss, EEG abnormalities, progressive atactic symptoms and cerebellar atrophy. The endocrine organs thyroid, liver and pancreas also appear affected. Patients show signs of hepatomegaly, pancreatic atrophy and hypothyroidism. (Picker-Minh *et al.*, 2016) Researchers are currently investigating therapeutic approaches using mTOR agonists in animal models. These approaches could help to slow down or even partially reverse Purkinje cell atrophy. (Picker-Minh *et al.*, 2023)

Our patients II.1 and II.2 also show all these previously described symptoms. All these similarities in their phenotype strongly suggest that the variants we identified in the *PTRH2* gene are disease-causing in this family. Further analyses should include the measurement of PTRH2 protein levels in fibroblasts.

Close monitoring of our patients is of utmost importance because this disease affects so many organ systems. There should be a close collaboration between physicians, neurologists, endocrinologists and other specialists to treat our patients as well as possible. Knowledge about the recessive inheritance also is essential for this family and their family planning.

4.1.3 *RNF216* – Family 58

In family 58 we identified two heterozygous variants in the *RNF216* gene. Within the family their segregation was confirmed.

RNF216 is the abbreviation for the Ring Finger Protein 216. It encodes the enzyme E3 ubiquitin ligase. This enzyme mediates ubiquitination and labelling of proteins for proteasome-mediated degradation suggesting that our found variants may disrupt this fundamental cellular process leading to a pathological impact on the cerebellum, hippocampus and cerebral white matter. Furthermore, there are pathological effects on the endocrine reproductive cascade due to damage in the hypothalamus and pituitary gland.

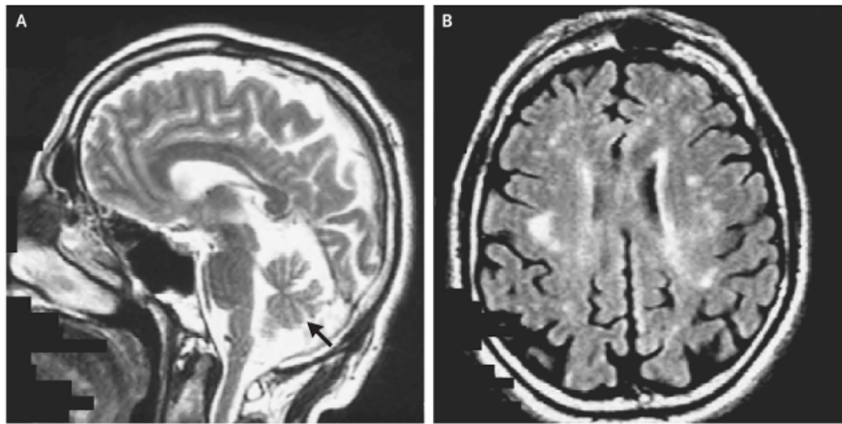

Figure 31: Panel A is a sagittal cerebral magnetic resonance image showing diffuse cerebellar atrophy (arrow) and cortical atrophy. Panel B is a transverse image and shows multiple distinct and confluent hyperintensity foci in the white matter. Reproduced with permission from "Ataxia, Dementia, and Hypogonadotropism Caused by Disordered Ubiquitination" by Margolin et al. 2013 (DOI: 10.1056/NEJMoa1215993), Copyright Massachusetts Medical Society.

Variants in this protein are already associated with autosomal recessive cerebellar ataxia and hypogonadotropic hypogonadism known as Gordon Holmes Syndrome. As shown in Figure 31 patients also present with cerebellar atrophy. In the animal model, which was demonstrated in zebrafish, inactivation of *RNF216* caused cerebellar disorganization and reduced the sizes of the head, optical tecta and eye cup. (Margolin *et al.*, 2013)

Our index patient takes Levetiracetam, testosterone and vitamin D per os. In addition, physiotherapeutic therapy is performed 1-2 times a week. Both drug and physiotherapeutic treatment should be continued.

4.1.4 *SBF1* – Family 59

In family 59 we identified two heterozygous variants in the *SBF1* gene.

SBF1 stands for Set-binding factor 1. It belongs to the myotubularin family and has a critical role regarding cellular growth control. (Cui *et al.*, 1998)

This gene is linked to the Charcot-Marie-Tooth disease which is a very heterogenous disease both clinically and genetically. (Patzko and Shy, 2012) It is a demyelinating disorder of the peripheral nervous system that is characterized by a slow progression with distal muscle weakness, atrophy and sensorimotor loss. Variants in the *SBF1* gene were linked to the Charcot-Marie-Tooth disease type 4B3 before. Patients with this disease reach early developmental milestones in time, but later present with progressive gait insecurities and weakness especially in their legs. There are signs of muscular atrophy of the upper and lower limbs, areflexia and foot deformities like a pes planus. (Nakhro *et al.*, 2013)

There are several clinical features that the patient shows that fit the phenotype. First symptoms were noticed when he was only 3 years old, so he achieved major motor milestones without any delay in his early development. The clinical examination showed gait insecurities with slow progression, absent Achilles tendon reflexes and distinctive hollow feet on both sides as well as reduced nerve conduction velocities. In addition to his physical impairments, he also has a low intelligence quotient and a short span of attention. The cMRI performed in 2013 showed no morphologically certain correlate for the stated symptoms, in particular there was no evidence of circumscribed atrophy or a barrier disorder.

As for now annual re-presentation in the neurological outpatient clinic as well as regular physical activity with coordination training as part of the physiotherapy is recommended.

4.1.5 *SEPSECS* – Family 81

In family 81 we identified a homozygous variant in the *SEPSECS* gene that segregates within the family.

The *SEPSECS* gene encodes for the enzyme O-phosphoserine-tRNA-selenocysteine-tRNA-synthase. Amino acids are chemical compounds consisting of an amino and a carboxylic acid group, which can be found in every living organism. They are the main components of proteins and are set free by proteolysis. The mRNA encodes the amino acid sequence in triplets with every triplet representing a codon that stands for a proteinogenic amino acid. All together they form a protein. There are 21 proteinogenic amino acids. Selenocysteine is the 21st amino acids and quite unique, because there is no codon representing selenocysteine in humans. (Heinrich, Müller and Graeve, 2014)

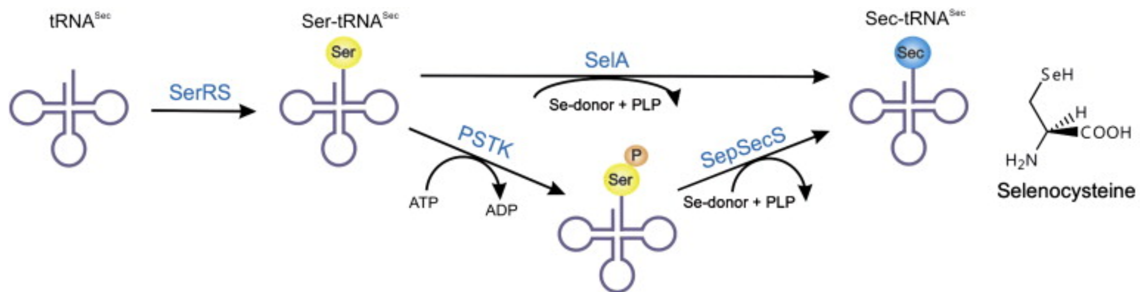

Figure 32: Biosynthesis of Sec-tRNA. The upper arrow shows the synthesis in bacteria using the selenocysteine synthase. The lower arrow shows the synthesis in eukaryotes and archae using an intermediate step and the enzyme SEPSECS to synthesize Sec-tRNA. Reprinted from The American Journal of Human Genetics, Agamy et al., Mutations Disrupting Selenocysteine Formation Cause Progressive Cerebello-Cerebral Atrophy, 2010, https://doi.org/10.1016/j.ajhg.2010.09.007, with permission from Elsevier

As shown in Figure 32 the formation of selenocysteine occurs on its own cognate tRNA and the enzyme O-phosphoserine-tRNA-selenocysteine-tRNA-synthase catalyses the final reaction in the selenocysteine synthesis by transforming O-phosphoseryl-tRNA into selenocysteinyl-tRNA. (Palioura *et al.*, 2009)

Progressive cerebello-cerebral atrophy is an autosomal recessive inherited disease that is linked to a missense variant in the *SEPSECS* gene. Patients with that disease present with mental retardation, microcephaly, spasticity and seizures. Imaging reveals a progressive cerebellar und cerebral atrophy. (Agamy *et al.*, 2010)

SEPSECS variants are also linked to the pontocerebellar hypoplasia type 2D. MRI findings in these patients seem similar to our patients' imaging results.

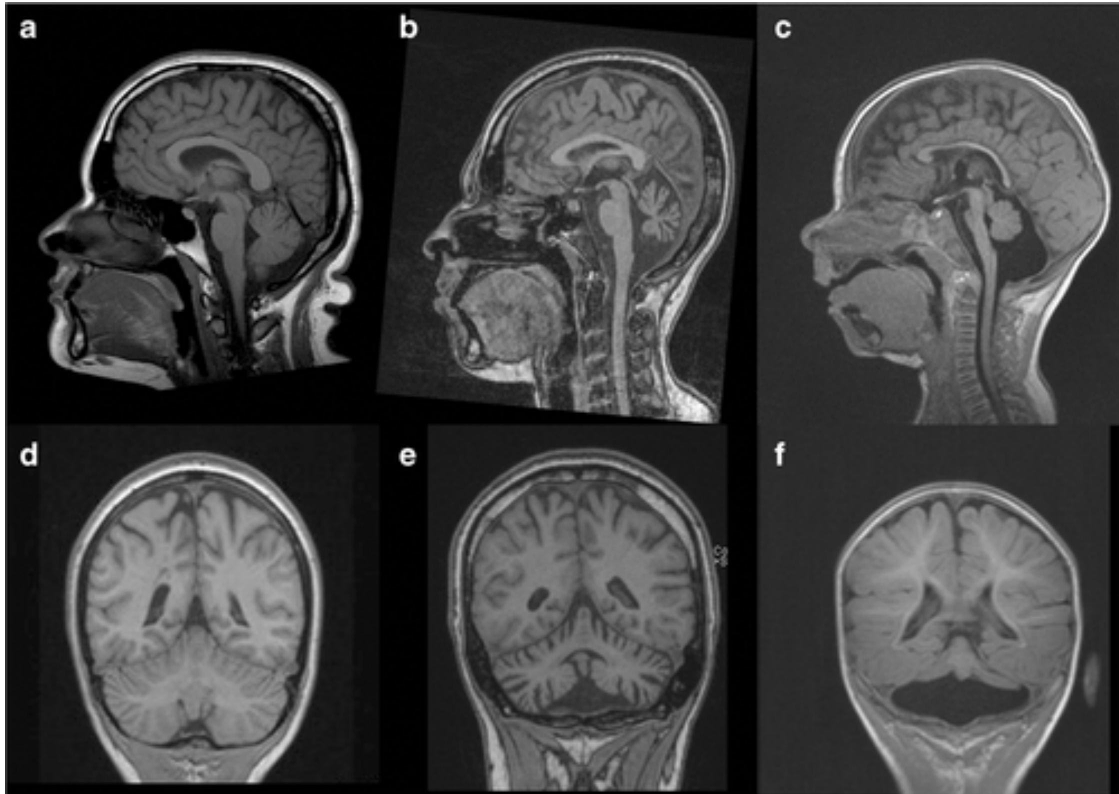

Figure 33: Cerebral magnetic resonance images of a patient with cerebellar atrophy (b), compared to a control (a) and a PCH2A patient (c). Source: van Dijk, T., Vermeij, JD., van Koningsbruggen, S. et al. A SEPSECS mutation in a 23-year-old woman with microcephaly and progressive cerebellar ataxia. J Inherit Metab Dis 41, 897–898 (2018). https://doi.org/10.1007/s10545-018-0151-x, http://creativecommons.org/licenses/by/4.0/

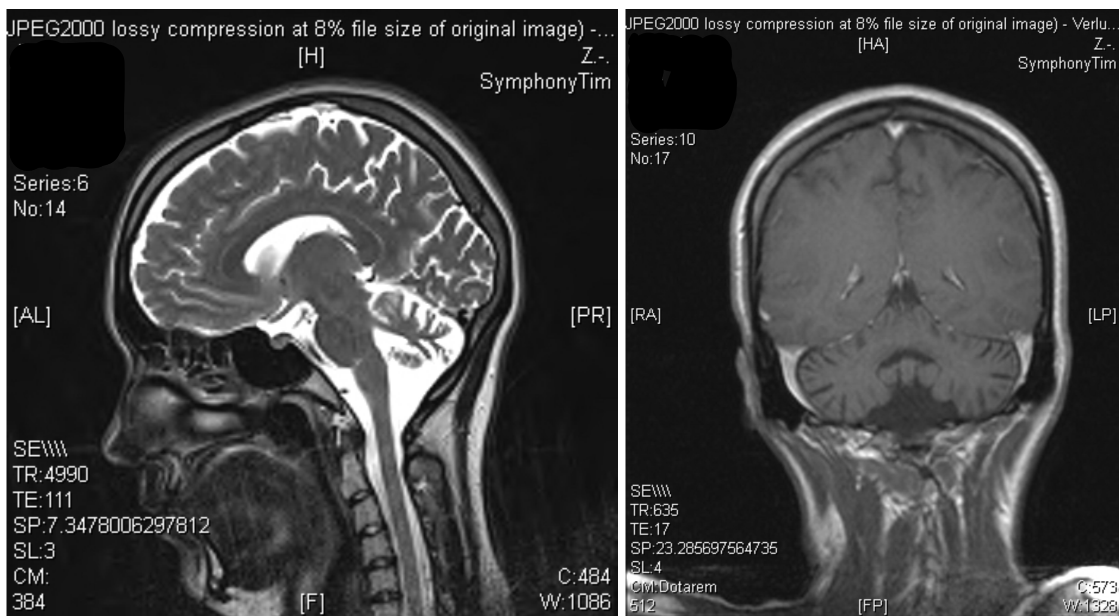

Figure 34: Sagittal (left) and coronar (right) cerebral magnetic resonance image of our index patient showing a significant atrophy of the cerebellum. There are no signs that would be indicative of ARSACS or leukoencephalopathy.

As shown in Figure 33 and Figure 34 they present with a hypoplastic cerebellum, small cerebellar hemispheres, a hypoplastic ventral pons, supratentorial cerebral atrophy, dilated ventricle, the white matter myelination appears delayed and there is no defect in the corpus callosum. (Uhl *et al.*, 1998; van Dijk *et al.*, 2018)

Further investigations to confirm the pathogenicity of this variant should also include the measurement of SepSecs activity by an anaerobic assay.

4.1.6 *NGLY1* – Family 88

In family 88 a homozygous variant was identified and its segregation within the family was confirmed via PCR and subsequent Sanger sequencing.

Transcription and translation of the highly conserved *NGLY1* gene lead to the enzyme N-glycanase that is very important for the protein degradation associated with the ER. Misfolded N-linked glycoproteins cannot be proceeded at the Golgi complex so the N-glycanase catalyzes their deglycosylation and therefore leads to their degradation. (Zhou *et al.*, 2006)

Variants in the *NGLY1* gene can lead to the congentinal disorder of deglycolysation 1 with partial or complete loss of function of the enzyme N-glycanase. Previous described patients with NGLY1-deficiency show a variety of phenotypes with developmental delay, mental retardation, microcephaly, epileptic seizures, movement disorders, elevated liver enzymes, scoliosis and muscular atrophy. (Enns *et al.*, 2014; Cahan and Frick, 2019) Due to the small number of patients, there is only a small amount of phenotypic information available.

In a clinical examination the index patient is described as friendly but agitated and restless. When she was one year old first hyperkinetic movement disorders were noticed. She would fell over when sitting and began toe walking when she was ten months old. At the age of five she was able to walk independently but lost this ability and was wheelchair-bound when turning 13 years old. She also has a combined speech expression disorder, dysphagia, an impairment of her fine motor skills and is mentally retarded. Her liver enzymes are slightly elevated. She has scoliosis and recurrent urinary tract infections as well as micturition disturbances. An electroencephalography showed epileptiforme potentials. Her affected siblings showed a very similar phenotype.

It is of keen interest to find a treatment for affected individuals. Using Drosophila there are now several compounds that were identified to potentially treat the NGLY deficiency. (Hope *et al.*, 2022; Pandey *et al.*, 2022)

AAV9-mediated gene replacement therapy in NGLY1-deficient rats normalises motor deficits and biomarkers in the brain. (Zhu *et al.*, 2022) A phase 1/2 study of the intracerebroventricular administration of AAV9-based GS-100 therapy is currently

underway in children and adolescents. Initial preclinical data show a reduction in neurological biomarkers, as well as an improvement in the behaviour of the test subjects.

4.1.7 *MTPAP* – Family 16

In family 16 we identified a homozygous synonymous variant in the *MTPAP* gene. *MTPAP* encodes for the mitochondrial poly (A) polymerase that catalyses the polyadenylation of the mRNA, which is important in many respects. It stabilizes the mRNA, leads to their nuclear export and its translation and also promotes the decay of the RNA. (Tomecki *et al.*, 2004) In human mitochondria the mitochondrial poly (A) polymerase is the key enzyme that implements the polyadenylation of the mRNA and therefore maintains the gene expression in the mitochondrion.

There is a known ataxia causing variant in the *MTPAP* gene that is inherited in an autosomal recessive manner. In an Amish family there are six children suffering from a progressive cerebellar ataxia, spastic paraparesis, dysarthria, optic atrophy and mental retardation. A homozygous c.1432A>G variant in the *MTPAP* gene was identified the segregates within the family. Further analyses were carried out like a poly (A) tail assay which showed a significant difference in length of the poly a tails leading to the conclusion that this variant is responsible for the spastic ataxia in this family. (Crosby *et al.*, 2010)

Mitochondrial poly (A) tail assays were performed to further investigate the pathogenic potential. Those assays showed a normal polyadenylation status leading to the conclusion that this variant does not cause the disease in this family.

4.2 Solved by other laboratories

There is an urgent need in research to not only follow one approach but to also collaborate with other research groups. Due to that three families could be solved by other laboratories.

4.2.1 *ATAD3A* – Family 26

In family 26 the Institute of Medical Genetics and Applied Genomics at the University Hospital Tübingen identified two variants in the *ATAD3A* gene segregating in the family.

There is evidence that variants in the *ATAD3A* gene can cause congenital cataract, cerebellar atrophy, hearing loss and chronic progressive external ophthalmoplegia. Using drosophila mutants, the pathogenicity and lethal potential of variants in the *ATAD3A* was shown. The above-mentioned p.Leu77Val variant was also replicated and compared to other variants. Drosophila flies with the p.Leu77Val variant showed a shorter lifespan and age-dependent motor deficits. Furthermore, flight analyses were performed, which were normal in those Drosophila flies at young age but showed deficits at old age. Compared to other variants, the p.Leu77Val variant seems to have less pathogenic potential on its own. (Yap *et al.*, 2021)

4.2.2 *MT_ATP6* – Family 39

MT_ATP6 gene encodes for the ATP synthase 6 which is also known as complex V of the respiratory chain. (Anderson *et al.*, 1981)

There are only few cases with affected individuals carrying a *MT_ATP6* variant. They are showing a broad spectrum of clinical symptoms. At the biochemical level, a lower ATP synthesis rate, an obtained ATP hydrolysis capacity and an abnormally higher mitochondrial membrane potential were found most frequently among variant carriers. Furthermore, their heteroplasmy level is much higher compaired to asymptomatic relatives. There is a positive correlation between heteroplasmy levels and phenotype expression, with higher heteroplasmy levels observed in earlier-onset phenotypes. (Ganetzky *et al.*, 2019)

4.2.3 *FXN* – Family 69

In family 69 a homoallelic GAA repeat expansion with about 1000 repeats was verified. *FXN* is the abbreviation for Frataxin. This is a mitochondrial protein and linked to Friedreich ataxia. (Schmucker *et al.*, 2008)

Friedreich ataxia, first noted in 1863, is a recessively inherited progressive and degenerative disease. The disease-causing variant is a homozygous trinucleotide repeat expansion located on chromosome 9. It leads to symptoms in the central and peripheral nervous systems and can also cause complication with heart, skeleton and endocrine pancreas. Depending on the length of the repeat expansion is the expression of the disease. In patients with shorter repeat expansions, a later onset of symptoms and a rather mild course is observed. The longer the repeat expansion is, the more severe the course of the disease. (Koeppen, 2011)

Due to the wide variety of symptoms, regular neurological follow-up is recommended, as well as regular consultations with cardiologists, orthopedists, family physicians and ophthalmologists.

Considerable therapeutic progress has been made in recent years in the treatment of Friedreich's ataxia. Omaveloxolone has been approved for symptomatic therapy of Friedreich ataxia since 2023. It activates the Nrf2 pathways, thereby reducing oxidative stress. (Lynch *et al.*, 2023)

Current causal treatment studies focus on epigenetic modulators, gene therapies and protein-based approaches. Despite promising data, robust long-term studies and combinatorial therapy concepts are still lacking in order to adequately address the heterogeneous and multisystemic nature of the disease.

4.3 New candidate genes

The aim in this study was to identify new ataxia-causing candidate genes. We identified 21 new potentially ataxia-causing genes in silico, shown in Table 29.

Following a thorough literature review, we prioritised further analysis of the GPAA1 variants identified. At the time of data collection, this gene had not yet been associated with ataxia. The other identified genes, which are listed in Table 29, have only been studied in silico thus far due to time constraints, as studying them experimentally would have exceeded the scope of our study. This could provide a basis for further research.

4.3.1 *GPAA1* – Families 43 and 62

We were able to identify two families that both carry variants in the *GPAA1* gene.

In family 43 and in family 62 we identified two homozygous missense variants in the *GPAA1* gene segregating within both families. Both variants were undescribed at the time of data collection in this study.

The *GPAA1* gene encodes for the glycosylphosphatidylinositol anchor attachment protein 1 and is an interactor with the *TCTN3* gene, which is an already known ataxia causing gene. The tissue expression is not specific as shown in Figure 35 and in Figure 36.

Bulk tissue gene expression for GPAA1 (ENSG00000197858.10)

Figure 35: Tissue gene expression of GPAA1. It shows low tissue specificity (source: https://gtexportal.org/home/gene/GPAA1).

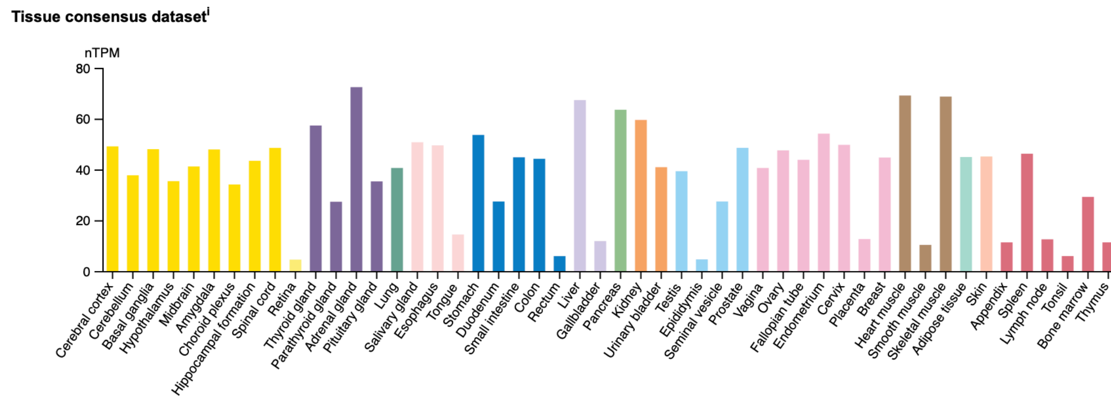

Figure 36: Tissue expression of GPAA1. There is no specific tissue expression (source: https://www.proteinatlas.org/ENSG00000197858-GPAA1).

First described in 1995 GPAA1 is an essential enzyme that enables the GPI-anchoring of proteins which will be attached to the surface of the cell membrane. (Hamburger, Egerton and Riezman, 1995)

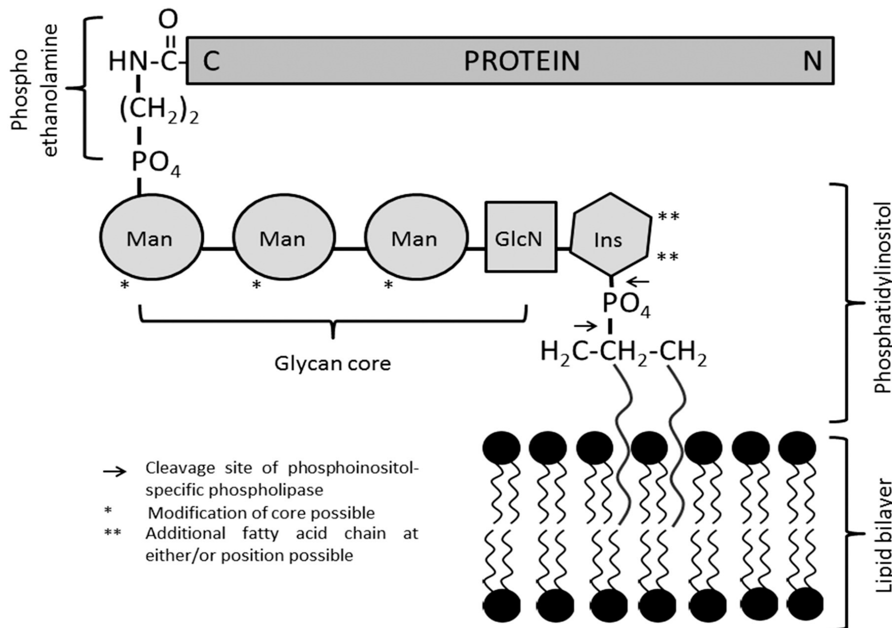

Figure 37: GPI anchored protein and its structural features. Reprinted from Journal of Lipid Research, Susanne Heider,John A. Dangerfield,Christoph Metzner et al., Biomedical applications of glycosylphosphatidylinositol-anchored proteins, 2016, https://doi.org/10.1194/jlr.R070201, http://creativecommons.org/licenses/by/4.0/

Proteins are covalently bound to GPI anchors via their carboxyl group as shown in Figure 37. GPI anchors are composed of ethanolamine phosphate, a trimannoside glycan core,

glucosamine and inositol phospholipid. (Heider, Dangerfield and Metzner, 2016) Their main function is to attach GPI-modified proteins at the cell surface. Therefore, after processing in the endoplasmatic reticulum, the GPI-modified protein is transferred as a lipid raft to the plasma membrane via trans-Golgi-network. (Ikezawa, 2002)

The amount of GPI anchored proteins on the cell surface gives us conclusions regarding the protein function. Quantification showed 30% less GPI anchored proteins leading to our conclusion that this might be an ataxia causing gene.

In 2017 Nguyen et al. also identified ten patients with *GPAA1* variants. Phenotypic features found were a delay in development, seizures, cerebellar atrophy and osteopenia. Functional studies showed reduced cell-surface quantities of GPI anchored proteins. (Nguyen *et al.*, 2017)

A further publication expands the range of symptoms associated with a variant in the *GPAA1* gene to include dysarthria, dysmetria and ataxia. (Castle *et al.*, 2021)

In 2023, another homozygous variant in the *GPAA1* gene was identified in a boy exhibiting similar symptoms to those observed in the two families in our cohort. He presented with hypotonia and strabismus and has a history of seizures. His cMRI scan shows progressive cerebellar atrophy. (Fontana *et al.*, 2023)

In order to investigate the effects on brain development and neurological function, Lukacs et al. developed a novel mouse model featuring targeted CNS-specific GPI anchor deficiency. Mice with GPI production switched off showed severe, progressive ataxia and cerebellar development defects. Life expectancy was shortened and the test animals died prematurely. These results highlight the importance of GPI anchors for normal brain development and neurological functioning. (Lukacs, Blizzard and Stottmann, 2020)

This supports our finding that variants in *GPAA1* interfering with the GPI anchor synthesis pathway could lead to the atactic symptoms in our patients in family 43 and 62.

There are also other variants in the GPI anchor synthesis leading to diseases like paroxysmal nocturnal hemoglobinuria, intellectual disability, hypotonia, epilepsy (Makrythanasis *et al.*, 2016) and mabry syndrome (Thompson *et al.*, 2012).

Further functional studies are needed to explain the variability in the patients' phenotypes.

For the families that have not been solved so far, further environmental diagnostics as well as a search based on a different inheritance pattern is recommended.

Furthermore, regular re-analyses are necessary because technological progress is increasingly leading to the discovery of variants that were previously unclear or unknown, which could result in a genetic diagnosis for these families.

5 Abstract

Hereditary cerebellar ataxias manifest themselves both clinically and genetically as extremely heterogeneous movement disorders. The disorders of balance and movement coordination that characterize these diseases are caused by a progressive degeneration of cerebellar Purkinje cells. Optic atrophy, retinitis pigmentosa, cognitive impairment, epileptic seizures and peripheral neuropathies are some of the other symptoms that can occur in connection with cerebellar ataxia.

To date, variants are known in 251 genes that can cause autosomal dominant, autosomal recessive, X-linked and mitochondrial forms of cerebellar ataxia. Despite this high number, about 50% of patients with ataxic symptoms do not have a genetically confirmed diagnosis.

The identification of the disease-causing genes is essential for the development of standardized, comprehensive and cost-effective genetic tests. On the other hand, it also forms the basis for therapy-orientated research and the development of effective and targeted therapies.

In this work, I am focusing on autosomal recessive inherited ataxias. I identified possible disease-causing gene alterations in silico and confirmed them by PCR and subsequent Sanger sequencing. In the cohort, which consists of 103 families, variants in known ataxia-causing genes were detected in 6 families. A synonymous variant in the *MTPAP* gene was classified as non-disease-causing by further investigation at the protein level.

In addition, I identified the *GPAA1* gene as a new ataxia-causing gene in 2 other families.

A thorough functional investigation of newly discovered disease-causing genes and variants is being conducted separately from this project.

6 Zusammenfassung

Hereditäre cerebelläre Ataxien manifestieren sich sowohl klinisch als auch genetisch als extrem heterogene Bewegungsstörungen. Die für diese Erkrankungen charakteristischen Störungen der Balance und der Bewegungskoordination entstehen aufgrund einer progressiven Degeneration cerebellärer Purkinje Zellen. Optikusatrophie, Retinitis pigmentosa, kognitive Beeinträchtigungen, epileptische Anfälle, periphere Neuropathien sind einige weitere Symptome, die im Zusammenhang mit einer cerebellären Ataxie auftreten können.

Bislang sind Varianten in ca. 250 Genen bekannt, die autosomal dominante, autosomal rezessive, X-chromosomale und mitochondriale Formen der cerebellären Ataxien verursachen können. Trotz dieser hohen Zahl haben in etwa 50% der Patienten mit ataktischen Symptomen keine genetisch gesicherte Diagnose.

Die Identifizierung der krankheitsverursachenden Gene ist zum einen für die Diagnostik zur Entwicklung standardisierter, umfangreicher und kostengünstiger Gentests essenziell. Zum anderen stellt sie aber auch die Basis für die therapieorientierte Forschung und Entwicklung wirksamer und zielgerichteter Therapien dar.

In meiner Doktorarbeit konzentriere ich mich auf autosomal rezessiv vererbte Ataxien. Ich identifizierte mögliche krankheitsverursachende Genveränderungen in silico und bestätige sie mittels PCR und anschließender Sanger-Sequenzierung. Insgesamt wurden in der Kohorte, die aus 103 Familien besteht, in 6 Familien Varianten in bekannten Ataxie-verursachenden Genen nachgewiesen. Eine synonyme Variante im *MTPAP* Gen wurde durch weiterführende Untersuchung auf Protein-Ebene als nicht krankheitsauslösend eingestuft.

Zudem habe ich das *GPAA1* Gen als neues, Ataxie-verursachendes Gen in 2 weiteren Familien identifiziert.

Eine gründliche funktionelle Untersuchung neu entdeckter krankheitsverursachender Gene und Varianten wird separat von diesem Projekt durchgeführt.

7 References

Agamy, O. *et al.* (2010) 'Mutations disrupting selenocysteine formation cause progressive cerebello-cerebral atrophy', *American Journal of Human Genetics*, 87(4), pp. 538–544. Available at: https://doi.org/10.1016/j.ajhg.2010.09.007.

Ahmadian, N. *et al.* (2019) 'The Cerebellar Cognitive Affective Syndrome—a Meta-analysis', *The Cerebellum*, 18(5), pp. 941–950. Available at: https://doi.org/10.1007/s12311-019-01060-2.

Akli, S. *et al.* (1990) 'A "G" to "A" mutation at position -1 of a 5' splice site in a late infantile form of Tay-Sachs disease.', *Journal of Biological Chemistry*, 265(13), pp. 7324–7330. Available at: https://doi.org/10.1016/S0021-9258(19)39117-3.

Amirifar, P. *et al.* (2019) 'Ataxia-telangiectasia: A review of clinical features and molecular pathology', *Pediatric Allergy and Immunology*, 30(3), pp. 277–288. Available at: https://doi.org/10.1111/pai.13020.

Anderson, S. *et al.* (1981) 'Sequence and organization of the human mitochondrial genome', *Nature*, 290(5806), pp. 457–465. Available at: https://doi.org/10.1038/290457a0.

Anheim, M. *et al.* (2010) 'Epidemiological, clinical, paraclinical and molecular study of a cohort of 102 patients affected with autosomal recessive progressive cerebellar ataxia from Alsace, Eastern France: implications for clinical management', *neurogenetics*, 11(1), pp. 1–12. Available at: https://doi.org/10.1007/s10048-009-0196-y.

Arslan, E.A. *et al.* (2017) 'Childhood hereditary ataxias: experience from a tertiary referral university hospital in Turkey', *Acta Neurologica Belgica*, 117(4), pp. 857–865. Available at: https://doi.org/10.1007/s13760-017-0786-7.

Baderna, V. *et al.* (2020) 'A novel AFG3L2 mutation close to AAA domain leads to aberrant OMA1 and OPA1 processing in a family with optic atrophy', *Acta Neuropathologica Communications*, 8(1), p. 93. Available at: https://doi.org/10.1186/s40478-020-00975-w.

Beal, M.F., Lang, A.E. and Ludolph, A.C. (2005) *Neurodegenerative Diseases: Neurobiology, Pathogenesis and Therapeutics*. Cambridge University Press.

Becker, E.B.E. *et al.* (2009) ‘A point mutation in TRPC3 causes abnormal Purkinje cell development and cerebellar ataxia in moonwalker mice’, *Proceedings of the National Academy of Sciences*, 106(16), pp. 6706–6711. Available at: https://doi.org/10.1073/pnas.0810599106.

Buckner, R.L. (2013) ‘The Cerebellum and Cognitive Function: 25 Years of Insight from Anatomy and Neuroimaging’, *Neuron*, 80(3), pp. 807–815. Available at: https://doi.org/10.1016/j.neuron.2013.10.044.

Cabal-Herrera, A.M. *et al.* (2020) ‘Fragile X-Associated Tremor/Ataxia Syndrome (FXTAS): Pathophysiology and Clinical Implications’, *International Journal of Molecular Sciences*, 21(12), p. 4391. Available at: https://doi.org/10.3390/ijms21124391.

Cahan, E.M. and Frick, S.L. (2019) ‘Orthopaedic phenotyping of NGLY1 deficiency using an international, family-led disease registry’, *Orphanet Journal of Rare Diseases*, 14(1), p. 148. Available at: https://doi.org/10.1186/s13023-019-1131-4.

Caporali, L. *et al.* (2020) ‘ATPase Domain AFG3L2 Mutations Alter OPA1 Processing and Cause Optic Neuropathy’, *Annals of Neurology*, 88(1), pp. 18–32. Available at: https://doi.org/10.1002/ana.25723.

Carson, A.R. *et al.* (2014) ‘Effective filtering strategies to improve data quality from population-based whole exome sequencing studies’, *BMC Bioinformatics*, 15(1), p. 125. Available at: https://doi.org/10.1186/1471-2105-15-125.

Castle, A.M.R. *et al.* (2021) ‘Expanding the Phenotypic Spectrum of GPI Anchoring Deficiency Due to Biallelic Variants in GPAA1’, *Neurology Genetics*, 7(6), p. e631. Available at: https://doi.org/10.1212/NXG.0000000000000631.

Crosby, A.H. *et al.* (2010) ‘Defective mitochondrial mRNA maturation is associated with spastic ataxia’, *American Journal of Human Genetics*, 87(5), pp. 655–660. Available at: https://doi.org/10.1016/j.ajhg.2010.09.013.

Cui, X. *et al.* (1998) ‘Association of SET domain and myotubularin-related proteins modulates growth control’, *Nature Genetics*, 18(4), pp. 331–337. Available at: https://doi.org/10.1038/ng0498-331.

Dietrichs, E. (2008) ‘Clinical manifestation of focal cerebellar disease as related to the organization of neural pathways’, *Acta Neurologica Scandinavica*, 117(s188), pp. 6–11. Available at: https://doi.org/10.1111/j.1600-0404.2008.01025.x.

van Dijk, T. *et al.* (2018) ‘A SEPSECS mutation in a 23-year-old woman with microcephaly and progressive cerebellar ataxia’, *Journal of Inherited Metabolic Disease*, 41(5), pp. 897–898. Available at: https://doi.org/10.1007/s10545-018-0151-x.

Durr, A. (2010) ‘Autosomal dominant cerebellar ataxias: polyglutamine expansions and beyond’, *The Lancet Neurology*, 9(9), pp. 885–894. Available at: https://doi.org/10.1016/S1474-4422(10)70183-6.

El-Hattab, A.W. *et al.* (2015) ‘MELAS syndrome: Clinical manifestations, pathogenesis, and treatment options’, *Molecular Genetics and Metabolism*, 116(1), pp. 4–12. Available at: https://doi.org/10.1016/j.ymgme.2015.06.004.

Enns, G.M. *et al.* (2014) ‘Mutations in NGLY1 cause an inherited disorder of the endoplasmic reticulum-associated degradation pathway’, *Genetics in Medicine: Official Journal of the American College of Medical Genetics*, 16(10), pp. 751–758. Available at: https://doi.org/10.1038/gim.2014.22.

Fogel, B.L. *et al.* (2014) ‘Exome Sequencing in the Clinical Diagnosis of Sporadic or Familial Cerebellar Ataxia’, *JAMA Neurology*, 71(10), pp. 1237–1246. Available at: https://doi.org/10.1001/jamaneurol.2014.1944.

Fontana, P. *et al.* (2023) ‘A Novel Homozygous GPAA1 Variant in a Patient with a Glycosylphosphatidylinositol Biosynthesis Defect’, *Genes*, 14(7), p. 1444. Available at: https://doi.org/10.3390/genes14071444.

Gaalen, J. van and Warrenburg, B.P.C. van de (2012) ‘A practical approach to late-onset cerebellar ataxia: putting the disorder with lack of order into order’, *Practical Neurology*, 12(1), pp. 14–24. Available at: https://doi.org/10.1136/practneurol-2011-000108.

Ganetzky, R.D. *et al.* (2019) ‘MT-ATP6 mitochondrial disease variants: Phenotypic and biochemical features analysis in 218 published cases and cohort of 14 new cases’, *Human Mutation*, 40(5), pp. 499–515. Available at: https://doi.org/10.1002/humu.23723.

Gillespie, J.H. (2004) *Population genetics: a concise guide*. 2nd ed. Baltimore, Md: Johns Hopkins University Press.

Gonzalez, M. *et al.* (2015) ‘Innovative genomic collaboration using the GENESIS (GEM.app) platform’, *Human mutation*, 36(10), pp. 950–956. Available at: https://doi.org/10.1002/humu.22836.

Gorman, G.S. *et al.* (2015) ‘Prevalence of nuclear and mitochondrial DNA mutations related to adult mitochondrial disease’, *Annals of Neurology*, 77(5), pp. 753–759. Available at: https://doi.org/10.1002/ana.24362.

Griffiths (2002) *Modern genetic analysis: Integrating genes and genomes.* 2nd edn. W. H. Freeman.

Hagerman, P.J. (2008) ‘The fragile X prevalence paradox’, *Journal of medical genetics*, 45(8), pp. 498–499. Available at: https://doi.org/10.1136/jmg.2008.059055.

Hamburger, D., Egerton, M. and Riezman, H. (1995) ‘Yeast Gaa1p is required for attachment of a completed GPI anchor onto proteins’, *The Journal of Cell Biology*, 129(3), pp. 629–639. Available at: https://doi.org/10.1083/jcb.129.3.629.

Heider, S., Dangerfield, J.A. and Metzner, C. (2016) ‘Biomedical applications of glycosylphosphatidylinositol-anchored proteins’, *Journal of Lipid Research*, 57(10), pp. 1778–1788. Available at: https://doi.org/10.1194/jlr.R070201.

Heinrich, P.C., Müller, M. and Graeve, L. (2014) *Löffler/Petrides Biochemie und Pathobiochemie*. Springer-Verlag.

Hersheson, J., Haworth, A. and Houlden, H. (2012) ‘The inherited ataxias: genetic heterogeneity, mutation databases, and future directions in research and clinical diagnostics’, *Human Mutation*, 33(9), pp. 1324–1332. Available at: https://doi.org/10.1002/humu.22132.

Hoche, F. *et al.* (2016) ‘Cerebellar Contribution to Social Cognition’, *The Cerebellum*, 15(6), pp. 732–743. Available at: https://doi.org/10.1007/s12311-015-0746-9.

Hope, K.A. *et al.* (2022) ‘An in vivo drug repurposing screen and transcriptional analyses reveals the serotonin pathway and GSK3 as major therapeutic targets for NGLY1

deficiency', *PLoS genetics*, 18(6), p. e1010228. Available at: https://doi.org/10.1371/journal.pgen.1010228.

Houge, G. *et al.* (2021) 'Stepwise ABC system for classification of any type of genetic variant', *European Journal of Human Genetics*, 30. Available at: https://doi.org/10.1038/s41431-021-00903-z.

Hu, H. *et al.* (2014) 'Mutations in PTRH2 cause novel infantile-onset multisystem disease with intellectual disability, microcephaly, progressive ataxia, and muscle weakness', *Annals of Clinical and Translational Neurology*, 1(12), pp. 1024–1035. Available at: https://doi.org/10.1002/acn3.149.

Huang, M. and Verbeek, D.S. (2019) 'Why do so many genetic insults lead to Purkinje Cell degeneration and spinocerebellar ataxia?', *Neuroscience Letters*, 688, pp. 49–57. Available at: https://doi.org/10.1016/j.neulet.2018.02.004.

Hwang, Y.-C. *et al.* (2015) 'HIPPIE: a high-throughput identification pipeline for promoter interacting enhancer elements', *Bioinformatics (Oxford, England)*, 31(8), pp. 1290–1292. Available at: https://doi.org/10.1093/bioinformatics/btu801.

Ikezawa, H. (2002) 'Glycosylphosphatidylinositol (GPI)-anchored proteins', *Biological & Pharmaceutical Bulletin*, 25(4), pp. 409–417. Available at: https://doi.org/10.1248/bpb.25.409.

Jacobi, H. and Minnerop, M. (2021) 'Ataxien des Erwachsenenalters', *Der Nervenarzt*, 92(4), pp. 379–389. Available at: https://doi.org/10.1007/s00115-021-01099-9.

Karczewski, K.J. *et al.* (2017) 'The ExAC browser: displaying reference data information from over 60 000 exomes', *Nucleic Acids Research*, 45(D1), pp. D840–D845. Available at: https://doi.org/10.1093/nar/gkw971.

Karczewski, K.J. *et al.* (2020) 'The mutational constraint spectrum quantified from variation in 141,456 humans', *Nature*, 581(7809), pp. 434–443. Available at: https://doi.org/10.1038/s41586-020-2308-7.

Kashiwabuchi, N. *et al.* (1995) 'Impairment of motor coordination, Purkinje cell synapse formation, and cerebellar long-term depression in GluRδ2 mutant mice', *Cell*, 81(2), pp. 245–252. Available at: https://doi.org/10.1016/0092-8674(95)90334-8.

Kasumu, A. and Bezprozvanny, I. (2012) ‘Deranged Calcium Signaling in Purkinje Cells and Pathogenesis in Spinocerebellar Ataxia 2 (SCA2) and Other Ataxias’, *The Cerebellum*, 11(3), pp. 630–639. Available at: https://doi.org/10.1007/s12311-010-0182-9.

Klockgether, T. (2010) ‘Sporadic ataxia with adult onset: classification and diagnostic criteria’, *The Lancet Neurology*, 9(1), pp. 94–104. Available at: https://doi.org/10.1016/S1474-4422(09)70305-9.

Koeppen, A.H. (2011) ‘Friedreich’s ataxia: pathology, pathogenesis, and molecular genetics’, *Journal of the Neurological Sciences*, 303(1–2), pp. 1–12. Available at: https://doi.org/10.1016/j.jns.2011.01.010.

Koppen, M. *et al.* (2007) ‘Variable and tissue-specific subunit composition of mitochondrial m-AAA protease complexes linked to hereditary spastic paraplegia’, *Molecular and Cellular Biology*, 27(2), pp. 758–767. Available at: https://doi.org/10.1128/MCB.01470-06.

Lecocq, C. *et al.* (2016) ‘Delayed-onset Friedreich’s ataxia revisited’, *Movement Disorders*, 31(1), pp. 62–69. Available at: https://doi.org/10.1002/mds.26382.

Lek, M. *et al.* (2016) ‘Analysis of protein-coding genetic variation in 60,706 humans’, *Nature*, 536(7616), pp. 285–291. Available at: https://doi.org/10.1038/nature19057.

Lim, J. *et al.* (2006) ‘A protein-protein interaction network for human inherited ataxias and disorders of Purkinje cell degeneration’, *Cell*, 125(4), pp. 801–814. Available at: https://doi.org/10.1016/j.cell.2006.03.032.

Lin, C.-Y.R. and Kuo, S.-H. (2023) ‘Ataxias: Hereditary, Acquired, and Reversible Etiologies’, *Seminars in neurology*, 43(1), pp. 48–64. Available at: https://doi.org/10.1055/s-0043-1763511.

Liss, B. and Kätzel, D. (2019) ‘Kleinhirn’, in R. Brandes, F. Lang, and R.F. Schmidt (eds) *Physiologie des Menschen: mit Pathophysiologie*. Berlin, Heidelberg: Springer, pp. 597–607. Available at: https://doi.org/10.1007/978-3-662-56468-4_46.

Lonsdale, J. *et al.* (2013) ‘The Genotype-Tissue Expression (GTEx) project’, *Nature Genetics*, 45(6), pp. 580–585. Available at: https://doi.org/10.1038/ng.2653.

Lopriore, P. *et al.* (2024) 'Primary Coenzyme Q10 Deficiency-Related Ataxias', *Journal of Clinical Medicine*, 13(8), p. 2391. Available at: https://doi.org/10.3390/jcm13082391.

Lukacs, M., Blizzard, L.E. and Stottmann, R.W. (2020) 'CNS glycosylphosphatidylinositol deficiency results in delayed white matter development, ataxia and premature death in a novel mouse model', *Human Molecular Genetics*, 29(7), pp. 1205–1217. Available at: https://doi.org/10.1093/hmg/ddaa046.

Lynch, D.R. *et al.* (2023) 'Efficacy of Omaveloxolone in Friedreich's Ataxia: Delayed-Start Analysis of the MOXIe Extension', *Movement Disorders*, 38(2), pp. 313–320. Available at: https://doi.org/10.1002/mds.29286.

Makrythanasis, P. *et al.* (2016) 'Pathogenic Variants in PIGG Cause Intellectual Disability with Seizures and Hypotonia', *American Journal of Human Genetics*, 98(4), pp. 615–626. Available at: https://doi.org/10.1016/j.ajhg.2016.02.007.

Maltecca, F. *et al.* (2015) 'Purkinje neuron Ca2+ influx reduction rescues ataxia in SCA28 model', *The Journal of Clinical Investigation*, 125(1), pp. 263–274. Available at: https://doi.org/10.1172/JCI74770.

Manto, M. and Mariën, P. (2015) 'Schmahmann's syndrome - identification of the third cornerstone of clinical ataxiology', *Cerebellum & Ataxias*, 2, p. 2. Available at: https://doi.org/10.1186/s40673-015-0023-1.

Manto, M. and Marmolino, D. (2009) 'Cerebellar ataxias', *Current Opinion in Neurology*, 22(4), pp. 419–429. Available at: https://doi.org/10.1097/WCO.0b013e32832b9897.

Margolin, D.H. *et al.* (2013) 'Ataxia, dementia, and hypogonadotropism caused by disordered ubiquitination', *The New England Journal of Medicine*, 368(21), pp. 1992–2003. Available at: https://doi.org/10.1056/NEJMoa1215993.

Matilla-Dueñas, A. *et al.* (2014) 'Consensus Paper: Pathological Mechanisms Underlying Neurodegeneration in Spinocerebellar Ataxias', *Cerebellum (London, England)*, 13(2), pp. 269–302. Available at: https://doi.org/10.1007/s12311-013-0539-y.

Metz, G. *et al.* (2013) ‘Rating disease progression of Friedreich’s ataxia by the International Cooperative Ataxia Rating Scale: analysis of a 603-patient database’, *Brain*, 136(1), pp. 259–268. Available at: https://doi.org/10.1093/brain/aws309.

Murken, J. (2006) *Taschenlehrbuch Humangenetik*. 7th edn. Stuttgart: Thieme.

Nachbauer, W., Eigentler, A. and Boesch, S. (2015) ‘Acquired ataxias: the clinical spectrum, diagnosis and management’, *Journal of Neurology*, 262(5), pp. 1385–1393. Available at: https://doi.org/10.1007/s00415-015-7685-8.

Nakhro, K. *et al.* (2013) ‘SET binding factor 1 (SBF1) mutation causes Charcot-Marie-Tooth disease type 4B3’, *Neurology*, 81(2), pp. 165–173. Available at: https://doi.org/10.1212/WNL.0b013e31829a3421.

Németh, A.H. *et al.* (2013) ‘Next generation sequencing for molecular diagnosis of neurological disorders using ataxias as a model’, *Brain*, 136(10), pp. 3106–3118. Available at: https://doi.org/10.1093/brain/awt236.

Ng, S.B. *et al.* (2009) ‘Targeted capture and massively parallel sequencing of 12 human exomes’, *Nature*, 461(7261), pp. 272–276. Available at: https://doi.org/10.1038/nature08250.

Ng, S.B. *et al.* (2010) ‘Exome sequencing identifies the cause of a mendelian disorder’, *Nature Genetics*, 42(1), pp. 30–35. Available at: https://doi.org/10.1038/ng.499.

Ngo, K.J. *et al.* (2020) ‘A diagnostic ceiling for exome sequencing in cerebellar ataxia and related neurological disorders’, *Human mutation*, 41(2), pp. 487–501. Available at: https://doi.org/10.1002/humu.23946.

Nguyen, T.T.M. *et al.* (2017) ‘Mutations in GPAA1, Encoding a GPI Transamidase Complex Protein, Cause Developmental Delay, Epilepsy, Cerebellar Atrophy, and Osteopenia’, *American Journal of Human Genetics*, 101(5), pp. 856–865. Available at: https://doi.org/10.1016/j.ajhg.2017.09.020.

Pagani, F., Raponi, M. and Baralle, F.E. (2005) ‘Synonymous mutations in CFTR exon 12 affect splicing and are not neutral in evolution’, *Proceedings of the National Academy of Sciences of the United States of America*, 102(18), pp. 6368–6372. Available at: https://doi.org/10.1073/pnas.0502288102.

Palau, F. and Espinós, C. (2006) ‘Autosomal recessive cerebellar ataxias’, *Orphanet Journal of Rare Diseases*, 1, p. 47. Available at: https://doi.org/10.1186/1750-1172-1-47.

Palioura, S. *et al.* (2009) ‘The human SepSecS-tRNASec complex reveals the mechanism of selenocysteine formation’, *Science (New York, N.Y.)*, 325(5938), pp. 321–325. Available at: https://doi.org/10.1126/science.1173755.

Pandey, A. *et al.* (2022) ‘NGLY1 Deficiency, a Congenital Disorder of Deglycosylation: From Disease Gene Function to Pathophysiology’, *Cells*, 11(7), p. 1155. Available at: https://doi.org/10.3390/cells11071155.

Park, S.Y. *et al.* (2021) ‘Anatomical Location of the Vestibulocerebellar Tract in the Healthy Human Brain: A Diffusion Tensor Imaging Study’, *Brain Sciences*, 11(2), p. 199. Available at: https://doi.org/10.3390/brainsci11020199.

Pascual-Castroviejo, I. *et al.* (1994) ‘Primary Degeneration of the Granular Layer of the Cerebellum. A Study of 14 Patients and Review of the Literature’, *Neuropediatrics*, 25(4), pp. 183–190. Available at: https://doi.org/10.1055/s-2008-1073020.

Patzko, A. and Shy, M.E. (2012) ‘Charcot-Marie-Tooth disease and related genetic neuropathies’, *Continuum (Minneapolis, Minn.)*, 18(1), pp. 39–59. Available at: https://doi.org/10.1212/01.CON.0000411567.34085.da.

Paul, M.S. and Limaiem, F. (2025) ‘Histology, Purkinje Cells’, in *StatPearls*. Treasure Island (FL): StatPearls Publishing. Available at: http://www.ncbi.nlm.nih.gov/books/NBK545154/ (Accessed: 29 May 2025).

Picker-Minh, S. *et al.* (2016) ‘Phenotype variability of infantile-onset multisystem neurologic, endocrine, and pancreatic disease IMNEPD’, *Orphanet Journal of Rare Diseases*, 11(1), p. 52. Available at: https://doi.org/10.1186/s13023-016-0433-z.

Picker-Minh, S. *et al.* (2023) ‘PTRH2 is Necessary for Purkinje Cell Differentiation and Survival and its Loss Recapitulates Progressive Cerebellar Atrophy and Ataxia Seen in IMNEPD Patients’, *The Cerebellum*, 22(6), pp. 1137–1151. Available at: https://doi.org/10.1007/s12311-022-01488-z.

Pilotto, F. *et al.* (2023) 'Early molecular layer interneuron hyperactivity triggers Purkinje neuron degeneration in SCA1', *Neuron*, 111(16), pp. 2523-2543.e10. Available at: https://doi.org/10.1016/j.neuron.2023.05.016.

Pilotto, F. and Saxena, S. (2018) 'Epidemiology of inherited cerebellar ataxias and challenges in clinical research', *Clinical and Translational Neuroscience*, 2(2), p. 2514183X18785258. Available at: https://doi.org/10.1177/2514183X18785258.

Pollard, K.S. *et al.* (2010) 'Detection of nonneutral substitution rates on mammalian phylogenies', *Genome Research*, 20(1), pp. 110–121. Available at: https://doi.org/10.1101/gr.097857.109.

Pontén, F., Jirström, K. and Uhlen, M. (2008) 'The Human Protein Atlas--a tool for pathology', *The Journal of Pathology*, 216(4), pp. 387–393. Available at: https://doi.org/10.1002/path.2440.

Rentzsch, P. *et al.* (2019) 'CADD: predicting the deleteriousness of variants throughout the human genome', *Nucleic Acids Research*, 47(Database issue), pp. D886–D894. Available at: https://doi.org/10.1093/nar/gky1016.

Richard, I. and Beckmann, J.S. (1995) 'How neutral are synonymous codon mutations?', *Nature Genetics*, 10(3), p. 259. Available at: https://doi.org/10.1038/ng0795-259.

Richards, S. *et al.* (2015) 'Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology', *Genetics in Medicine: Official Journal of the American College of Medical Genetics*, 17(5), pp. 405–424. Available at: https://doi.org/10.1038/gim.2015.30.

Richter, U. *et al.* (2019) 'Mitochondrial stress response triggered by defects in protein synthesis quality control', *Life Science Alliance*, 2(1). Available at: https://doi.org/10.26508/lsa.201800219.

Ruano, L. *et al.* (2014) 'The global epidemiology of hereditary ataxia and spastic paraplegia: a systematic review of prevalence studies', *Neuroepidemiology*, 42(3), pp. 174–183. Available at: https://doi.org/10.1159/000358801.

Rudolph, S. *et al.* (2023) ‘Cognitive-Affective Functions of the Cerebellum’, *Journal of Neuroscience*, 43(45), pp. 7554–7564. Available at: https://doi.org/10.1523/JNEUROSCI.1451-23.2023.

Sakaguchi, S. *et al.* (1996) ‘Loss of cerebellar Purkinje cells in aged mice homozygous for a disrupted PrP gene’, *Nature*, 380(6574), pp. 528–531. Available at: https://doi.org/10.1038/380528a0.

Samocha, K. *et al.* (2017) *Regional missense constraint improves variant deleteriousness prediction*. Available at: https://doi.org/10.1101/148353.

Samocha, K.E. *et al.* (2014) ‘A framework for the interpretation of de novo mutation in human disease’, *Nature Genetics*, 46(9), pp. 944–950. Available at: https://doi.org/10.1038/ng.3050.

Sato, A. *et al.* (1999) ‘Adenovirus-Mediated Expression of Mutant DRPLA Proteins with Expanded Polyglutamine Stretches in Neuronally Differentiated PC12 Cells. Preferential Intranuclear Aggregate Formation and Apoptosis’, *Human Molecular Genetics*, 8(6), pp. 997–1006. Available at: https://doi.org/10.1093/hmg/8.6.997.

Schmahmann, J.D. (2004) ‘Disorders of the Cerebellum: Ataxia, Dysmetria of Thought, and the Cerebellar Cognitive Affective Syndrome’, *The Journal of Neuropsychiatry and Clinical Neurosciences*, 16(3), pp. 367–378. Available at: https://doi.org/10.1176/jnp.16.3.367.

Schmahmann, J.D. and Sherman, J.C. (1998) ‘The cerebellar cognitive affective syndrome.’, *Brain*, 121(4), pp. 561–579. Available at: https://doi.org/10.1093/brain/121.4.561.

Schmitz-Hübsch, T. *et al.* (2006) ‘Scale for the assessment and rating of ataxia: development of a new clinical scale’, *Neurology*, 66(11), pp. 1717–1720. Available at: https://doi.org/10.1212/01.wnl.0000219042.60538.92.

Schmucker, S. *et al.* (2008) ‘The in vivo mitochondrial two-step maturation of human frataxin’, *Human Molecular Genetics*, 17(22), pp. 3521–3531. Available at: https://doi.org/10.1093/hmg/ddn244.

Serra, H.G. *et al.* (2006) 'RORα-Mediated Purkinje Cell Development Determines Disease Severity in Adult *SCA1* Mice', *Cell*, 127(4), pp. 697–708. Available at: https://doi.org/10.1016/j.cell.2006.09.036.

Siepel, A. *et al.* (2005) 'Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes', *Genome Research*, 15(8), pp. 1034–1050. Available at: https://doi.org/10.1101/gr.3715005.

Simon, D. *et al.* (2004) 'Friedreich Ataxia Mouse Models with Progressive Cerebellar and Sensory Ataxia Reveal Autophagic Neurodegeneration in Dorsal Root Ganglia', *Journal of Neuroscience*, 24(8), pp. 1987–1995. Available at: https://doi.org/10.1523/JNEUROSCI.4549-03.2004.

Smeets, C.J.L.M. and Verbeek, D.S. (2014) 'Cerebellar ataxia and functional genomics: Identifying the routes to cerebellar neurodegeneration', *Biochimica Et Biophysica Acta*, 1842(10), pp. 2030–2038. Available at: https://doi.org/10.1016/j.bbadis.2014.04.004.

Synofzik, M., Schöls, L. and Riess, O. (2013) 'Hereditäre Ataxien', *medizinische genetik*, 25(2), pp. 235–248. Available at: https://doi.org/10.1007/s11825-013-0383-y.

Taroni, F., Chiapparini, L. and Mariotti, C. (2013) 'Autosomal Dominant Spinocerebellar Ataxias and Episodic Ataxias', in *Handbook of the Cerebellum and Cerebellar Disorders*. Springer, Dordrecht, pp. 2193–2267. Available at: https://doi.org/10.1007/978-94-007-1333-8_101.

Tavtigian, S.V. *et al.* (2020) 'Fitting a naturally scaled point system to the ACMG/AMP variant classification guidelines', *Human Mutation*, 41(10), pp. 1734–1737. Available at: https://doi.org/10.1002/humu.24088.

Taylor, A.M.R. and Byrd, P.J. (2005) 'Molecular pathology of ataxia telangiectasia', *Journal of Clinical Pathology*, 58(10), pp. 1009–1015. Available at: https://doi.org/10.1136/jcp.2005.026062.

Thapa, S. *et al.* (2022) 'Ataxia due to vitamin E deficiency: A case report and updated review', *Clinical Case Reports*, 10(9), p. e6303. Available at: https://doi.org/10.1002/ccr3.6303.

‘The Genotype-Tissue Expression (GTEx) project’ (2013) *Nature genetics*, 45(6), pp. 580–585. Available at: https://doi.org/10.1038/ng.2653.

Thompson, M.D. *et al.* (2012) ‘Phenotypic variability in hyperphosphatasia with seizures and neurologic deficit (Mabry syndrome)’, *American Journal of Medical Genetics. Part A*, 158A(3), pp. 553–558. Available at: https://doi.org/10.1002/ajmg.a.35202.

Tomecki, R. *et al.* (2004) ‘Identification of a novel human nuclear-encoded mitochondrial poly(A) polymerase’, *Nucleic Acids Research*, 32(20), pp. 6001–6014. Available at: https://doi.org/10.1093/nar/gkh923.

Trouillas, P. *et al.* (1997) ‘International Cooperative Ataxia Rating Scale for pharmacological assessment of the cerebellar syndrome. The Ataxia Neuropharmacology Committee of the World Federation of Neurology’, *Journal of the Neurological Sciences*, 145(2), pp. 205–211. Available at: https://doi.org/10.1016/s0022-510x(96)00231-6.

Uhl, M. *et al.* (1998) ‘MR findings in pontocerebellar hypoplasia’, *Pediatric Radiology*, 28(7), pp. 547–551. Available at: https://doi.org/10.1007/s002470050410.

Uhlén, M. *et al.* (2015) ‘Proteomics. Tissue-based map of the human proteome’, *Science (New York, N.Y.)*, 347(6220), p. 1260419. Available at: https://doi.org/10.1126/science.1260419.

Van Overwalle, F. *et al.* (2020) ‘Consensus Paper: Cerebellum and Social Cognition’, *The Cerebellum*, 19(6), pp. 833–868. Available at: https://doi.org/10.1007/s12311-020-01155-1.

Vedolin, L. *et al.* (2013) ‘Inherited cerebellar ataxia in childhood: a pattern-recognition approach using brain MRI’, *AJNR. American journal of neuroradiology*, 34(5), pp. 925–934, S1-2. Available at: https://doi.org/10.3174/ajnr.A3055.

Vernon, H.J. and Bindoff, L.A. (2018) ‘Chapter 9 - Mitochondrial ataxias’, in M. Manto and T.A.G.M. Huisman (eds) *Handbook of Clinical Neurology*. Elsevier (The Cerebellum: Disorders and Treatment), pp. 129–141. Available at: https://doi.org/10.1016/B978-0-444-64189-2.00009-3.

Winser, S.J. *et al.* (2018) ‘Can pre-screening vestibulocerebellar involvement followed by targeted training improve the outcomes of balance in cerebellar ataxia?’, *Medical Hypotheses*, 117, pp. 37–41. Available at: https://doi.org/10.1016/j.mehy.2018.06.001.

Wong, Y.Y. *et al.* (2010) ‘Cordycepin inhibits protein synthesis and cell adhesion through effects on signal transduction’, *The Journal of Biological Chemistry*, 285(4), pp. 2610–2621. Available at: https://doi.org/10.1074/jbc.M109.071159.

Yang, Z. (1995) ‘A space-time process model for the evolution of DNA sequences.’, *Genetics*, 139(2), pp. 993–1005. Available at: https://doi.org/10.1093/genetics/139.2.993.

Yap, Z.Y. *et al.* (2021) ‘Functional interpretation of ATAD3A variants in neuro-mitochondrial phenotypes’, *Genome Medicine*, 13(1), p. 55. Available at: https://doi.org/10.1186/s13073-021-00873-3.

Zhou, X. *et al.* (2006) ‘Structural and biochemical studies of the C-terminal domain of mouse peptide-N-glycanase identify it as a mannose-binding module’, *Proceedings of the National Academy of Sciences of the United States of America*, 103(46), pp. 17214–17219. Available at: https://doi.org/10.1073/pnas.0602954103.

Zhu, L. *et al.* (2022) ‘AAV9-NGLY1 gene replacement therapy improves phenotypic and biomarker endpoints in a rat model of NGLY1 Deficiency’, *Molecular Therapy. Methods & Clinical Development*, 27, pp. 259–271. Available at: https://doi.org/10.1016/j.omtm.2022.09.015.

Zühlke, C. *et al.* (2002) ‘Spinocerebellar ataxia type 1 (SCA1): Phenotype-genotype correlation studies in intermediate alleles’, *European Journal of Human Genetics*, 10(3), pp. 204–209. Available at: https://doi.org/10.1038/sj.ejhg.5200788.

8 Erklärungen zum Eigenanteil der Dissertationsschrift

Die Arbeit wurde in der Universitätsklinik Tübingen und dem Hertie-Institut für klinische Hirnforschung unter Betreuung von Frau Prof. Dr. med. Rebecca Schüle-Freyer durchgeführt.

Die Konzeption der Studie erfolgte durch Frau Prof. Dr. med. Rebecca Schüle-Freyer.

Sämtliche Versuche wurden nach Einarbeitung durch Frau Prof. Dr. med. Rebecca Schüle-Freyer und Frau Jennifer Reichbauer von mir teils eigenständig, teils mit Unterstützung durch Frau Jennifer Reichbauer durchgeführt.

Ich versichere, das Manuskript selbständig verfasst zu haben und keine weiteren als die von mir angegebenen Quellen verwendet zu haben.

Hennigsdorf, den

Anika Koch

9 Acknowledgements

First, I would like to thank Prof. Dr. Rebecca Schüle-Freyer for supporting me over this long time. Thank you for letting me be a part of your research group. I learned so much from you and I am still fascinated that you could always answer every single one of my questions immediately and had a solution for every problem.

I would also like to thank Prof. Dr. Matthis Synfozik. You were a great support and I could always ask you for your advice. Thank you for giving me great input for this work.

A big thank you also goes to Jennifer Reichbauer and Selina Reich. Thank you for training and supporting me in the lab and always having an open ear.

I would like to give the biggest thanks to my family. You mean everything to me. You have always supported me, believed in me and made this work possible in the first place. I love you.