ARTICLE OPEN ACCESS

Genetic background of ataxia in children younger than 5 years in Finland

Erika Ignatius, MD, Pirjo Isohanni, MD, PhD, Max Pohjanpelto, Päivi Lahermo, PhD, Simo Ojanen, MSc, Virginia Brilhante, PhD, Eino Palin, MD, PhD, Anu Suomalainen, MD, PhD, Tuula Lönnqvist, MD, PhD, and Christopher J. Carroll, PhD

Correspondence
Dr. Ignatius
erika.ignatius@helsinki.fi
or Dr. Carroll
ccarroll@sgul.ac.uk

Neurol Genet 2020;6:e444. doi:10.1212/NXG.0000000000000444

Abstract

Objective
To characterize the genetic background of molecularly undefined childhood-onset ataxias in Finland.

Methods
This study examined a cohort of patients from 50 families with onset of an ataxia syndrome before the age of 5 years collected from a single tertiary center, drawing on the advantages offered by next generation sequencing. A genome-wide genotyping array (Illumina Infinium Global Screening Array MD-24 v.2.0) was used to search for copy number variation undetectable by exome sequencing.

Results
Exome sequencing led to a molecular diagnosis for 20 probands (40%). In the 23 patients examined with a genome-wide genotyping array, 2 additional diagnoses were made. A considerable proportion of probands with a molecular diagnosis had de novo pathogenic variants (45%). In addition, the study identified a de novo variant in a gene not previously linked to ataxia: MED23. Patients in the cohort had medically actionable findings.

Conclusions
There is a high heterogeneity of causative mutations in this cohort despite the defined age at onset, phenotypical overlap between patients, the founder effect, and genetic isolation in the Finnish population. The findings reflect the heterogeneous genetic background of ataxia seen worldwide and the substantial contribution of de novo variants underlying childhood ataxia.

From the Department of Child Neurology (E.I., P.I., T.L.), Children's Hospital, University of Helsinki and Helsinki University Hospital; Research Programs Unit, Stem Cells and Metabolism, Faculty of Medicine (E.I., P.I., M.P., S.O., V.B., E.P., A.S.), Institute for Molecular Medicine Finland (FIMM) (P.L.), Neuroscience Center (A.S.), HiLife, University of Helsinki, Finland; and Genetics Research Centre (C.J.C.), Molecular and Clinical Sciences Research Institute, St. George's, University of London, United Kingdom.

Go to Neurology.org/NG for full disclosures. Funding information is provided at the end of the article.

The Article Processing Charge was funded by the Helsinki University Hospital.

GLOSSARY

ACMG = American College of Medical Genetics and Genomics; **CADD** = Combined Annotation Dependent Depletion; **CGH** = comparative genomic hybridization; **CNV** = copy number variant; **ExAC** = Exome Aggregation Consortium; **GSAMD** = Global Screening Array Multi-disease; **HPO** = Human Phenotype Ontology; **NGS** = next generation sequencing; **SNP** = single nucleotide polymorphism.

The most common etiology of ataxia in the pediatric population is genetic, with the prevalence of genetic childhood ataxia in Europe estimated at 14.6 per 100,000 population.[1] Determining the etiology of childhood-onset ataxia has important clinical relevance, including ending the stressful and costly diagnostic odyssey, guiding genetic counseling, and facilitating precise follow-up and treatment.

The most common causes of hereditary ataxia vary regionally in populations of different genetic backgrounds.[1] Owing to the founder effect and genetic isolation, Finland has a unique disease heritage.[2] Accordingly, the most common ataxias seen elsewhere in the world, such as Friedreich ataxia, are rare in Finland.

As next generation sequencing (NGS) technologies have evolved, there have been many reports of exome sequencing in single families or single cases with childhood-onset cerebellar ataxia. Many previously reported ataxia cohorts analyze patients with adult or varied age-of-onset[3] or are defined by having structural cerebellar abnormalities[4–6] instead of the symptom of ataxia. There are a few smaller studies that analyze cohorts of pediatric patients with the symptom of ataxia.[7,8]

This study applied exome sequencing and a genome-wide genotyping array to examine a cohort of patients with childhood-onset ataxia collected from a single tertiary center, allowing better characterization of the genetic background of molecularly undefined childhood-onset ataxias in Finland. An age limit of 5 years at onset was applied to demarcate a clinical entity within the heterogeneous group of hereditary ataxias and is based on the Human Phenotype Ontology (HPO)[9] defining childhood-onset as onset before the fifth birthday.

Methods

Patients

We reviewed all pediatric patients who were diagnosed with ataxia as the primary symptom of their disease in a single tertiary center (Helsinki University Hospital Child Neurology) during years 1999–2018. An overview of the study is shown in figure 1. The exclusion criteria are provided in appendix e-1 (links.lww.com/NXG/A269).

Thirty-three children from 25 families seen in the Helsinki University Hospital Child Neurology clinic between 1999 and 2018 for the onset of ataxia younger than 5 years had a molecular diagnosis made outside of this study. The etiology included pathogenic variants in the genes *ATXN7* (2 families), *CACNA1A* (4 families), *C12orf65* (2 families), *DNAJC19*, *FOLR1* (2 families), *NARP*, *NKX2-1*, *PDHA1*, *SCN2A*, *SCN8A*, *SLC17A5* (3 families), *SUCLA2*, and *TWNK* (4 families). In addition, a 6.4 Mb deletion in chromosome 10 (10q26.2q26.3) was found to underlie ataxia in 1 patient.

We invited families with children who received a clinical diagnosis of an ataxia syndrome of an unknown etiology and with onset of symptoms before the age of 5 years to participate in the study. We recruited 50 families while 2 families declined participation. In families with multiple affected children, we selected the first child referred to our clinic as the proband. From October 2014 through February 2019, we performed exome sequencing on samples obtained from 50 probands.

Routine genetic screening was not required before entering the study, although single gene or gene-panel testing of known ataxia genes had been performed for some patients before the study. In this regard, if the treating physician suspected a specific genetic etiology, for example, infantile-onset spinocerebellar ataxia, a disease belonging to the Finnish disease heritage, the gene in question may have been tested before recruitment into this study. Table 1 describes the clinical characteristics and demographics of the cohort. For all patients in the cohort, diagnostic investigations for ataxia had begun during early childhood. The results of a clinical microarray comparative genomic hybridization (CGH) test performed before the study were available for 46% (23/50) of all probands. Twelve patients (24%) were either prescreened or screened during the study for pathologically expanded trinucleotide repeats (appendix e-1, links.lww.com/NXG/A269).

Standard protocol approvals, registrations, and patient consents

The study was approved by the Helsinki University Hospital ethics review board. All patients and/or their legal guardians gave informed consent in accordance with the Declaration of Helsinki.

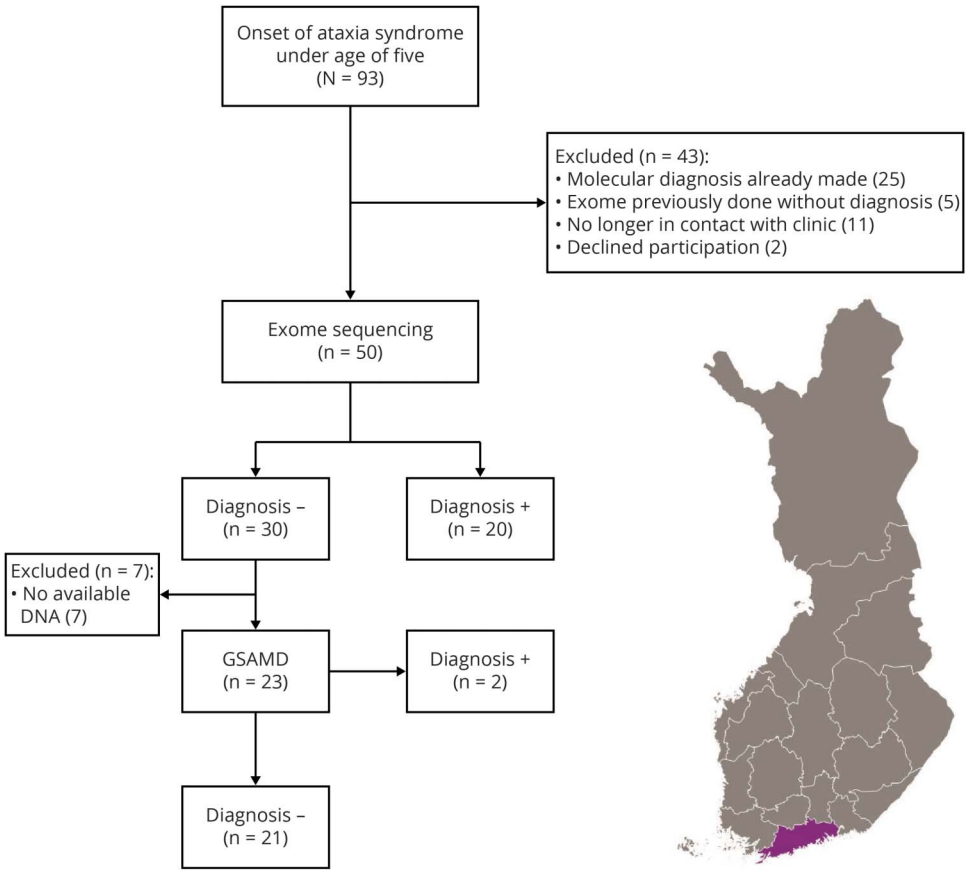
Phenotyping

A child neurologist with expertise in childhood-onset ataxia examined all patients, whereas all available clinical, laboratory, and imaging data were reviewed by several clinicians undertaking the study. Primary phenotypes were mapped to HPO terms[9] and included in the in-house semiautomated variant prioritization pipeline.

Sequencing and bioinformatics analysis

We performed exome sequencing on genomic DNA for 50 probands, 2 affected parents, and an affected sibling in 2 families.

Figure 1 Study flowchart



Number of families with children with ataxia syndrome seen in the Helsinki University Hospital Child Neurology Clinic in 1999–2018. Patients were referred from the Helsinki-Uusimaa region (catchment area population ~1.7 million, ~30% of the Finnish population). The Helsinki-Uusimaa region is highlighted on the map. GSAMD = Global Screening Array Multi-disease.

The variant calling pipeline of the Finnish Institute of Molecular Medicine was used for the reference genome alignment and variant calling.[10] We prioritized recessive-type nonsynonymous variants with a minor allele frequency of less than 0.1% on the Exome Aggregation Consortium (ExAC)[11] server. For potential de novo or dominantly inherited variants, heterozygous variants that were not found at all on the ExAC server were prioritized for consideration. We further prioritized variants by their predicted deleterious effect using amino acid conservation and, in the case of potential de novo variants, by gene constraint to mutation according to the framework previously described.[11] The prediction tools SIFT[12] and Polyphen[13] as well as Combined Annotation Dependent Depletion (CADD) C-score[14] were used in variant evaluation. Variants with a CADD C-score[14] of less than 10 were excluded. Variants were also compared with our in-house database containing 520 exomes.

We classified novel sequence variants using the guidelines provided by the American College of Medical Genetics and Genomics (ACMG).[15]

Sanger sequencing was used to validate the variants identified by exome sequencing and for segregation analysis. In the case of P12, samples from the child's biological parents were unavailable, and as confirmation, western blot was used to confirm the deleterious effect of the variant identified.

Technical information on exome sequencing, a list of the primers used in Sanger sequencing and details of experimental validation of variants identified for P3 and P12, is provided in appendix e-1 (links.lww.com/NXG/A269).

Global screening array analysis

We screened for copy number variation and uniparental disomy after exome analysis was negative in 23 probands, for whom there was available DNA, using the 759993 single nucleotide polymorphism (SNP) markers of Illumina Infinium Global Screening Array MD-24 v.2.0 (GSAMD; Illumina, San Diego, CA). Log R ratio and B allele frequency values were generated with GenomeStudio 2.0 software (Illumina), and copy number variation regions were detected with PennCNV software using standard quality control checks.[16] Standard quality control of genome-wide genotyping data was performed with PLINK 1.9 software.[17]

Data availability

The data that support the findings of this study are available on request. The data are not publicly available because of the information that could compromise the privacy of research participants.

Table 1 Demographic and clinical background of the cohort

Sex	n (%); n = 50
Male	25 (50.0)
Female	25 (50.0)
Ethnicity	n (%); n = 50
Finnish	40 (80.0)
Other Caucasian	3 (6.0)
Middle East	2 (4.0)
Southern Asia	2 (4.0)
Multiple populations	2 (4.0)
Unknown	1 (2.0)
Pedigree	n (%); n = 50
Nonconsanguineous	
One affected child	39 (78.0)
Two or more affected children	5 (10.0)
Consanguineous	
One affected child	2 (4.0)
Two or more affected children	—
Affected parent	
One affected child	2 (4.0)
Two or more affected children	2 (4.0)
Age at onset of first symptoms	n (%); n = 50
Congenital (at birth)	3 (6.0)
Neonatal (birth–28 d)	5 (10.0)
Infantile (28 d–1 y)	26 (52.0)
Childhood (1–5 y)	16 (32.0)
Imaging abnormality	n (%); n = 50
Presence of abnormality on brain MRI	33 (66.0)
Cerebellar atrophy	22 (44.0)
Cerebellar atrophy and additional findings[a]	3 (6.0)
Cerebellar hypoplasia or dysplasia	4 (8.0)
Cerebellar hypoplasia	1 (2.0)
Cerebellar dysplasia	1 (2.0)
Hypoplasia/aplasia of the vermis and molar tooth sign	2 (4.0)
Other abnormality[b]	4 (8.0)
Abnormality of peripheral nerve conduction	n (%); n = 17
Presence of abnormality on ENMG	5 (29.4)
Sensory axonal neuropathy	2 (11.8)
Sensorimotor axonal neuropathy	1 (5.9)
Demyelinating peripheral neuropathy	2 (11.8)
Intellectual developmental disability	n (%); n = 50
Presence of intellectual developmental disability	22 (44.0)

[a] Findings additional to cerebellar atrophy: brainstem atrophy (2); basal ganglia abnormality (1).
[b] Other abnormality: hypoplasia of the cerebellum, pons, brainstem and corpus callosum (1); pontocerebellar hypoplasia (1); small volume of the thalami (1); mild cerebral cortical atrophy (1).
Results of an ENMG study were available for 17 patients.

Results

Diagnostic yield of exome sequencing

We obtained a molecular diagnosis for 20 probands (40%) using exome sequencing. We identified 26 diagnostic variants in 16 genes, 13 of which were novel and include a variant in the gene *GPAA1* that was published from this cohort.[18] A recessive form of inheritance was found in 9 probands, with 3 of the diagnosed probands having homozygous variants and 6 having compound heterozygous variants. We identified dominant variants in 11 probands, including de novo variants in 8 probands, 2 familial autosomal dominant variants, and a suspicion of parental somatic mosaicism in 1 parent of 1 proband. In all familial cases, there were no pedigrees involving 3 or more generations.

After multidisciplinary evaluation, a novel variant in *COQ8A* and *STUB1* was considered diagnostic for the probands' ataxia although they remained as variants of uncertain significance as per strict application of the ACMG guidelines.

The diagnostic variants we found are listed in table 2. We annotated variants to the Ensembl[19] canonical transcript for each gene. Allele frequencies are reported as found in the Genome Aggregation Database.[20]

Diagnostic yield of genome-wide genotyping array

We uncovered 2 copy number variants (CNVs) that we considered pathogenic using the GSAMD genome-wide genotyping array. For P21, we identified a heterozygous 570 kb deletion in the 10q26.3 region (chr10:131538728-132108832), encompassing the genes *EBF3*, *GLRX*, *LINC00959*, and part of *MGMT*. The same deletion had already been identified in the patient's clinical molecular karyotype when analyzed in 2015 but was, at that time, considered to be of uncertain significance. The parents of the patient were screened for the mutation, and it was found to be de novo. The GSAMD finding prompted re-evaluation of the deletion. Since 2017, haploinsufficiency of *EBF3* has been reported to cause hypotonia, ataxia, and delayed development syndrome (MIM #617330).

For P22, we identified a heterozygous 1.2 kb deletion in *SLC2A1* (chr1:43392250-43393465), encompassing most of

Table 2 Diagnostic variants underlying the patients' ataxia

	Chr	Pos. Start	Gene	CDS position	Functional change	gnomAD AF Finnish	gnomAD AF All	CADD C-score	Inheritance	Reference/ClinVar ID if reported previously (annotation in report)	ACMG classification
P1	1	43394882	*SLC2A1*	c.971C>A	p.(Ser324Ter)	0	0	44	De novo	Novel variant	Pathogenic
P2	1	227169808	*COQ8A*	c.811C>T	p.(Arg271Cys)	0.0007748	0.0001061	33	AR	Reported[29]	Established variant
P2	1	227174171	*COQ8A*	c.1677C>G	p.(His559Gln)	0.002802	0.0004296	26.4	AR	ClinVar ID 423260	VUS
P3	2	191116992	*HIBCH*	c.559C>T	p.(Leu187Phe)	0	0.00003185	29	AR	Novel variant	Likely pathogenic[a]
P3	2	191159358	*HIBCH*	c.220-2A>T	r.spl	0	0.00001769	24.3	AR	Novel variant	Pathogenic
P4	2	219525942	*BCS1L*	c.232A>G	p.(Ser78Gly)	0.004062	0.0004737	23	AR	Reported[30]	Established variant
P4	2	219526485	*BCS1L*	c.464G>A	p.(Arg155Gln)	0.000008838	0.000003994	25.8	AR	Novel variant	Likely pathogenic
P5	3	4709128	*ITPR1*	c.1736C>T	p.(Thr579Ile)	0	0	29.5	AD	Reported[31] (Thr594Ile)	Established variant
P6	3	4856866	*ITPR1*	c.7786_7788delAAG	p.(Lys2596del)	0	0	22.8	De novo	Reported[32]	Established variant
P7	8	145138112	*GPAA1*	c.160_161delinsAA	p.(Ala54Asn)	0	0	29.9	AR	Reported from this cohort[18]	Established variant
P7	8	145139371	*GPAA1*	c.869T>C	p.(Leu290Pro)	0.00004704	0.000004020	26.1	AR	Reported from this cohort[18]	Established variant
P8	10	131640542	*EBF3*	c.1183C>T	p.(Arg395Ter)	0	0	42	AD[b]	ClinVar ID 620273	Pathogenic
P9	10	131676043	*EBF3*	c.625C>T	p.(Arg209Trp)	0	0	35	De novo	Reported[33]	Established variant
P10	10	131676045	*EBF3*	c.622dupA	p.(Met208AsnfsTer56)	0	0	35	AD	Novel variant	Pathogenic
P11	11	6636680	*TPP1*	c.1259C>A	p.(Ser420Ter)	0	0	40	AR	Novel variant	Likely pathogenic
P12	12	111057639	*TCTN1*	c.221-2A>G	r.spl	0	0	24.9	AR	reported[34] (IVS1–2A>G)	Established variant
P12	12	111085142	*TCTN1*	c.1635+1G>A	r.spl	0.000008845	0.000008022	24.8	AR	Novel variant	Pathogenic[a]
P13	13	77575055	*CLN5*	c.1175_1176delAT	p.(Tyr392Ter)	0.0008137	0.00009384	35	AR	Reported[35]	Established variant
P14	14	36987093	*NKX2-1*	c.596C>A	p.(Ser199Ter)	0	0	39	De novo	Reported[36] (C2519A)	Established variant
P15	16	732184	*STUB1*	c.689_692delACCT	p.(Tyr230CysfsTer9)	0.000008841	0.000007985	35	AR	Novel variant	Likely pathogenic
P15	16	732442	*STUB1*	c.865G>A	p.(Val289Ile)	0	0.000007998	23	AR	Novel variant	VUS
P16	17	57775212	*PTRH2*	c.127_128insA	p.(Ser43LysfsTer11)	0	0.000003979	24.3	AR	Novel variant	Likely pathogenic
P17	19	13346507	*CACNA1A*	c.4988G>A	p.(Arg1663Gln)	0	0	25.4	De novo	Reported[37] (Arg1664Gln)	Established variant

Continued

Table 2 Diagnostic variants underlying the patients' ataxia *(continued)*

	Chr	Pos. Start	Gene	CDS position	Functional change	gnomAD AF Finnish	gnomAD AF All	CADD C-score	Inheritance	Reference/ClinVar ID if reported previously (annotation in report)	ACMG classification
P18	19	42474691	*ATP1A3*	c.2306G>A	p.(Arg769His)	0	0	34	De novo	Reported[38] (Arg756His)	Established variant
P19	19	42474692	*ATP1A3*	c.2305C>T	p.(Arg769Cys)	0	0	34	De novo	Reported[39] (Arg756Cys)	Established variant
P20	X	41495865	*CASK*	c.879_880dupC	p.(Gln294ArgfsTer3)	0	0	35	De novo	Novel variant	Pathogenic

Abbreviations: AD = autosomal dominant; AF = allele frequency; AR = autosomal recessive; CADD = Combined Annotation Dependent Depletion; CDS = coding sequence; gnomAD = Genome Aggregation Database; VUS = variant of uncertain significance.
Variants are annotated to the Ensembl GRCh37 release 100 canonical transcript.
[a] Details of functional testing are in appendix e-1 (links.lww.com/NXG/A269).
[b] Maternal somatic mosaicism based on Sanger sequencing, mother clinically unaffected (appendix e-1, links.lww.com/NXG/A269).

exon 9 and all of exon 10. Mutations in *SLC2A1* are known to cause GLUT1 deficiency syndrome 1 (MIM #606777) and GLUT1 deficiency syndrome 2 (MIM #612126). Multiexon deletions in *SLC2A1* are known to cause disease.[21] The GSAMD finding was confirmed with multiplex ligation-dependent probe amplification in a clinical laboratory and confirmed to be de novo.

Patient phenotypes and effect on clinical management

The findings of the study affected the clinical management of multiple patients. Table e-1 (links.lww.com/NXG/A268) describes the patient phenotypes and possible effects on clinical treatment and management.

Variants of uncertain significance in ataxia genes

Four probands had a variant of uncertain significance in a gene previously implicated in ataxia, listed in table 3.

Gene of uncertain significance

We found a heterozygous de novo missense variant for P27 in a gene not previously linked to ataxia, *MED23* (figure 2A). The patient has hypotonia, tremor, and ataxia that developed at the age of 1.5 years. A detailed phenotypic description for P27 is in appendix e-1 (links.lww.com/NXG/A269). *MED23* encodes a transcription factor in which recessive mutations are known to cause autosomal recessive nonsyndromic mental retardation-18 (MIM #614249).[22] De novo status was confirmed by DNA fingerprinting of the patient and parents using 7 microsatellite markers. *MED23* has a high constraint for missense mutations (missense Z: 4.53556). The variant, chr6g.131919485A>G, c.2549T>C, p.(Leu850Pro) has a high CADD C-score (28.6) and causes the change of a conserved amino acid (figure 2B).

Discussion

Exome sequencing is a robust diagnostic method for childhood-onset ataxias manifesting before the age of 5. We found disease-causing mutations in many different genes in this cohort despite the defined age at onset, phenotypical overlap between patients, the founder effect, and genetic isolation in the Finnish population. This is surprising because Finland has a unique disease heritage; however, our findings reflect the heterogeneous genetic background of ataxia seen worldwide and the substantial contribution of de novo variants underlying childhood ataxia.

The patients who were investigated with exome sequencing formed a "hard-to-diagnose" cohort, which did not include patients from the same clinic whose genetic diagnosis had previously been made by single gene or panel testing. In this study, the combination of exome sequencing and GSAMD provided a diagnosis for 44% of the investigated families. This is slightly higher than in ataxia cohorts comprising patients with varying ages of disease onset, where an estimated diagnostic rate for exome sequencing is 36%.[3] Our diagnostic yield was at

Table 3 Variants of uncertain significance for the probands' ataxia

Proband	Chr	Pos. Start	Gene	CDS position	Functional change	CADD C-score	GnomAD AF Finnish	GnomAD AF All	Inheritance	Phenotype	Comment
P23	2	166245891	*SCN2A*	c.5575T>C	p.(Phe1859Leu)	24.9	0	0	AD	Episodic ataxia, brain MRI normal, and childhood-onset disease	Affected parent and sibling carry variant
P24	3	4687363	*ITPR1*	c.806G>T	p.(Arg269Leu)	34	0	0	AD	Ataxia, intellectual disability, cerebellar atrophy, and infantile-onset disease	Inherited from the affected parent
P25	12	52080889	*SCN8A*	c.500G>A	p.(Gly167Glu)	32	0	0	AD	Ataxia, epilepsy, intellectual disability, widened subarachnoid spaces, and infantile onset disease	Inherited from the unaffected parent
P26	18	6950834	*LAMA1*	c.8344 G>A	p.(Gly2782Ser)	28.5	0	0	AR	Ataxia, epilepsy, intellectual disability, brain MRI normal, and childhood-onset disease	
P26	18	6999503	*LAMA1*	c.4604C>T	p.(Ser1535Leu)	23.2	0.0006387	0.00007799	AR	Ataxia, epilepsy, intellectual disability, brain MRI normal, and childhood-onset disease	

Abbreviations: AD = autosomal dominant; AF = allele frequency; AR = autosomal recessive; CADD = Combined Annotation Dependent Depletion; CDS = coding sequence; gnomAD = Genome Aggregation Database. Variants are annotated to the Ensembl GRCh37 release 100 canonical transcript. Onset: infantile (28 days–1 year), childhood (1–5 years).

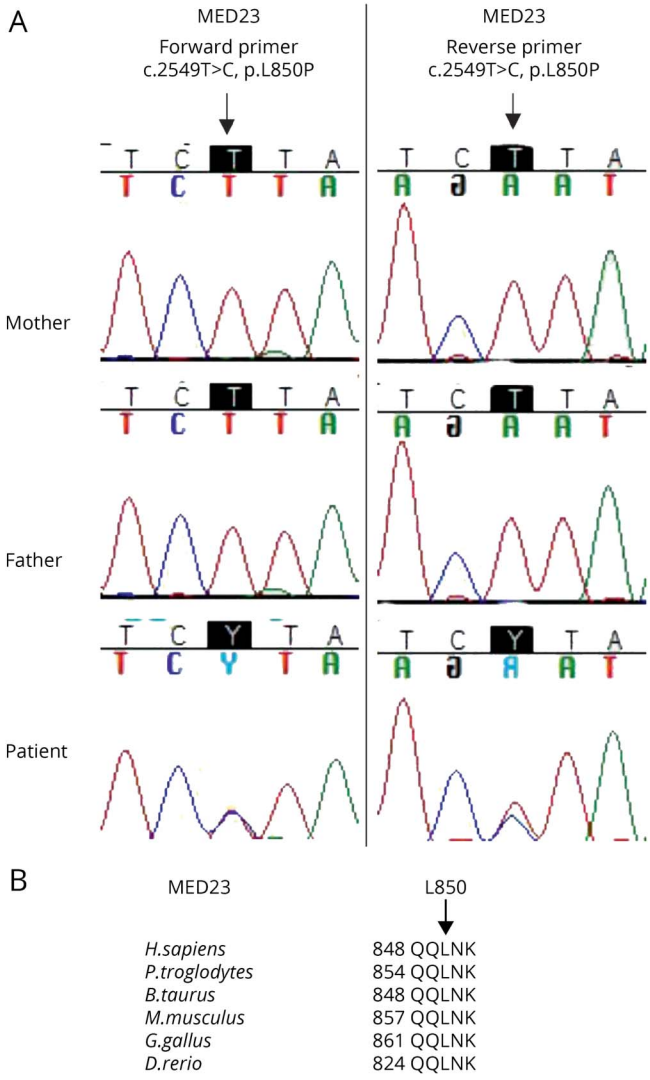
the same level as other smaller exome sequencing studies of symptom-based pediatric ataxia (46%–80%).[7,8,23]

The percentage of genetically diagnosed patients that were found to have a de novo variant underlying their disease in previous studies varied. In the first such published study,[7] in a cohort of 28 families (6 consanguineous), 9% had a pathogenic de novo variant. Our study revealed a remarkably higher de novo rate of 45% in patients with previously molecularly undiagnosed childhood-onset ataxia. Other reports include a 25% de novo rate in a study that investigated congenital ataxia in consanguineous families[8] and another study[23] that found a 42% de novo rate in a pediatric movement disorder cohort including patients with ataxia.

One of the limitations in our study was that in most cases only the proband, and not parents, was sequenced. Trio exome analysis is associated with a higher diagnostic yield compared with single exome analysis in rare Mendelian disorders,[24] and trio analysis is especially useful in an early onset, mainly sporadic ataxia cohort.[3] However, a large number of de novo mutations were still identified by prioritizing deleterious heterozygous variants using CADD C-scores and gene constraint scores.[11,14] Nevertheless, the high burden of de novo variants in this cohort adds to the recommendation of a triosequencing approach. Furthermore, our exome sequencing analysis may have underestimated the number of CNVs and may have overlooked uniparental disomy because their analysis is not straightforward from exome sequencing data. Software that infer copy number variation from exome sequencing data, such as ExomeDepth,[25] can have suboptimal specificity and sensitivity, especially in the case of small CNVs spanning one exon.[26] In children with developmental delay, a first-tier diagnostic test revealing CNVs has been chromosomal microarray, either with array-based CGH or SNP array.[27] Not all copy number variation is revealed even with the combination of microarray CGH and exome sequencing because CNVs at the intragene level will usually be too small to be identified using the microarrays in clinical use. High-resolution microarray or genome sequencing can still better identify CNVs, especially the long indels and small CNVs that are otherwise not found. We found GSAMD to be a cost-effective method to screen for copy number variation and uniparental disomy and to confirm maternity and paternity in suspected de novo cases in a research setting. In the patients examined with GSAMD, 4 had previous findings reported in clinical molecular karyotypes. These previously identified variations had been inherited from an unaffected parent or had otherwise been interpreted to be of uncertain significance. The GSAMD detected all 4 of these CNVs, adding to our confidence in the method.

We did not find any pathologic repeat expansions, such as those underlying Friedreich ataxia or many of the dominant spinocerebellar ataxias, which are poorly detected using NGS technologies. Many of the common spinocerebellar ataxias caused by trinucleotide repeats are rare in Finland, with the exception of spinocerebellar ataxia type 7,[28] and are unlikely to be represented in an early-childhood-onset cohort. However, repeat expansions

Figure 2 MED23 variant



The P27 heterozygous variant, chr6g.131919485A>G, c.2549T>C, p.(Leu850-Pro) in *MED23* (A) causes the change of a conserved amino acid (B).

in novel ataxia disease genes are possibly yet to be discovered, and identification of such mutations is likely to be enabled by application of long-read genome sequencing technologies.

NGS techniques have become ubiquitous diagnostic methods in centers studying neurodevelopmental and neurodegenerative disorders. Developing effective diagnostic algorithms requires experience of the utility of these methods as first- or second-line studies. As new genes and broader phenotypes in ataxia continue to be identified, targeted gene panels may overlook recently identified disease genes. In the case of our cohort, many findings would not have been made using the panels available at the time of sequencing. In the case of a negative exome or genome, systematic re-evaluation at a later time point may reveal a diagnosis. Publications of candidate genes potentially causing diseases in humans, as well as internet resources listing rare variants, aid researchers in finding other families with the same disease.

Most Finnish patients with childhood-onset ataxia are candidates for exome or genome sequencing when the phenotype and background do not clearly point to a specific disease entity. Patients in the cohort had medically actionable findings, underscoring the importance of exome, or genome sequencing as a first-line diagnostic method. Concurrently, it is crucial for the clinician to understand the inherent weaknesses of exome sequencing, especially the inefficiency concerning current analysis tools to detect copy number variation and triplet repeats.

Acknowledgment

The authors thank all the patients and relatives that participated in this study as well as Milla Kuronen, Lilli Hinds, Tuula Manninen, Markus Innilä, and Anu Harju for technical assistance.

Study funding

This work was funded by the University of Helsinki, Helsinki University Hospital, Arvo and Lea Ylppö Foundation, Foundation for Pediatric Research, Päivikki and Sakari Sohlberg Foundation, Biomedicum Helsinki Foundation and Maire Taponen Foundation.

Disclosure

The authors report no disclosures relevant to the manuscript. Go to Neurology.org/NG for full disclosures.

Publication history

Received by *Neurology: Genetics* January 3, 2020. Accepted in final form April 27, 2020.

Appendix Authors

Name	Location	Contribution
Erika Ignatius, MD	Children's Hospital, University of Helsinki and Helsinki University Hospital; and Research Programs Unit, University of Helsinki, Finland	Research project conception, organization, and execution; acquisition, analysis, and interpretation of data; and manuscript: writing of the first draft and review and critique
Pirjo Isohanni; MD, PhD	Children's Hospital, University of Helsinki and Helsinki University Hospital; and Research Programs Unit, University of Helsinki, Finland	Research project conception, organization, and execution; acquisition, analysis, and interpretation of data; and manuscript: review and critique
Max Pohjanpelto	Research Programs Unit, University of Helsinki, Finland	Research project execution; acquisition, analysis, and interpretation of data; and manuscript: review and critique
Päivi Lahermo, PhD	Institute for Molecular Medicine Finland, University of Helsinki, Finland	Research project execution; acquisition, analysis, and interpretation of data; and manuscript: writing of the first draft and review and critique

Appendix *(continued)*

Name	Location	Contribution
Simo Ojanen, MSc	Research Programs Unit, University of Helsinki, Finland	Research project organization; analysis and interpretation of data; and manuscript: review and critique
Virginia Brilhante, PhD	Research Programs Unit, University of Helsinki, Finland	Research project organization; analysis and interpretation of data; and manuscript: review and critique
Eino Palin, MD, PhD	Research Programs Unit, University of Helsinki, Finland	Research project organization; analysis and interpretation of data; and manuscript: review and critique
Anu Suomalainen, MD, PhD	Research Programs Unit, University of Helsinki, Finland	Research project conception and organization; and manuscript: review and critique
Tuula Lönnqvist, MD, PhD	Children's Hospital, University of Helsinki and Helsinki University Hospital, Finland	Research project conception, organization, and execution; acquisition, analysis, and interpretation of data; and manuscript: review and critique
Christopher J. Carroll, PhD	St. George's, University of London, United Kingdom	Research project conception, organization, and execution; acquisition, analysis, and interpretation of data; and manuscript: review and critique

References

1. Musselman KE, Stoyanov CT, Marasigan R, et al. Prevalence of ataxia in children: a systematic review. Neurology 2014;82:80–89.
2. Norio R. The Finnish Disease Heritage III: the individual diseases. Hum Genet 2003; 112:470–526.
3. Galatolo D, Tessa A, Filla A, Santorelli FM. Clinical application of next generation sequencing in hereditary spinocerebellar ataxia: increasing the diagnostic yield and broadening the ataxia-spasticity spectrum. A retrospective analysis. Neurogenetics 2018;19:1–8.
4. Chemin J, Siquier-Pernet K, Nicouleau M, et al. De novo mutation screening in childhood-onset cerebellar atrophy identifies gain-of-function mutations in the CACNA1G calcium channel gene. Brain 2018;141:1998–2013.
5. Megahed H, Nicouleau M, Barcia G, et al. Utility of whole exome sequencing for the early diagnosis of pediatric-onset cerebellar atrophy associated with developmental delay in an inbred population. Orphanet J Rare Dis 2016;11:57.
6. Ohba C, Osaka H, Iai M, et al. Diagnostic utility of whole exome sequencing in patients showing cerebellar and/or vermis atrophy in childhood. Neurogenetics 2013;14:225–232.
7. Sawyer SL, Schwartzentruber J, Beaulieu CL, et al. Exome sequencing as a diagnostic tool for pediatric-onset ataxia. Hum Mutat 2014;35:45–49.
8. Valence S, Cochet E, Rougeot C, et al. Exome sequencing in congenital ataxia identifies two new candidate genes and highlights a pathophysiological link between some congenital ataxias and early infantile epileptic encephalopathies. Genet Med 2019;21:553–563.
9. Kohler S, Vasilevsky NA, Engelstad M, et al. The Human Phenotype Ontology in 2017. Nucleic Acids Res 2017;45:D865–D876.
10. Sulonen AM, Ellonen P, Almusa H, et al. Comparison of solution-based exome capture methods for next generation sequencing. Genome Biol 2011;12:R94.
11. Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature 2016;536:285–291.
12. Sim NL, Kumar P, Hu J, Henikoff S, Schneider G, Ng PC. SIFT web server: predicting effects of amino acid substitutions on proteins. Nucleic Acids Res 2012;40:W452–W457.
13. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. Nat Methods 2010;7:248–249.
14. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet 2014;46:310–315.
15. Richards S, Aziz N, Bale S, et al. 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. Genet Med 2015;17:405–424.
16. Wang K, Li M, Hadley D, et al. PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. Genome Res 2007;17:1665–1674.
17. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559–575.
18. Nguyen TTM, Murakami Y, Sheridan E, et al. Mutations in GPAA1, encoding a GPI transamidase complex protein, cause developmental delay, epilepsy, cerebellar atrophy, and osteopenia. Am J Hum Genet 2017;101:856–865.
19. Zerbino DR, Achuthan P, Akanni W, et al. Ensembl 2018. Nucleic Acids Res 2018;46: D754–D761.
20. Karczewski KJ, Francioli LC, Tiao G, et al. Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. bioRxiv 2019:531210.
21. Leen WG, Klepper J, Verbeek MM, et al. Glucose transporter-1 deficiency syndrome: the expanding clinical and genetic spectrum of a treatable disorder. Brain 2010;133:655–670.
22. Hashimoto S, Boissel S, Zarhrate M, et al. MED23 mutation links intellectual disability to dysregulation of immediate early gene expression. Science 2011;333:1161–1163.
23. Cordeiro D, Bullivant G, Siriwardena K, et al. Genetic landscape of pediatric movement disorders and management implications. Neurol Genet 2018;4:e265.
24. Lee H, Deignan JL, Dorrani N, et al. Clinical exome sequencing for genetic identification of rare Mendelian disorders. JAMA 2014;312:1880–1887.
25. Plagnol V, Curtis J, Epstein M, et al. A robust model for read count data in exome sequencing experiments and implications for copy number variant calling. Bioinformatics 2012;28:2747–2754.
26. Samarakoon PS, Sorte HS, Kristiansen BE, et al. Identification of copy number variants from exome sequence data. BMC Genomics 2014;15:661.
27. Miller DT, Adam MP, Aradhya S, et al. Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. Am J Hum Genet 2010;86:749–764.
28. Jonasson J, Juvonen V, Sistonen P, et al. Evidence for a common Spinocerebellar ataxia type 7 (SCA7) founder mutation in Scandinavia. Eur J Hum Genet 2000;8:918–922.
29. Horvath R, Czermin B, Gulati S, et al. Adult-onset cerebellar ataxia due to mutations in CABC1/ADCK3. J Neurol Neurosurg Psychiatry 2012;83:174–178.
30. Visapää I, Fellman V, Vesa J, et al. GRACILE syndrome, a lethal metabolic disorder with iron overload, is caused by a point mutation in BCS1L. Am J Hum Genet 2002;71:863–876.
31. Sasaki M, Ohba C, Iai M, et al. Sporadic infantile-onset spinocerebellar ataxia caused by missense mutations of the inositol 1,4,5-triphosphate receptor type 1 gene. J Neurol 2015;262:1278–1284.
32. McEntagart M, Williamson KA, Rainger JK, et al. A restricted repertoire of de novo mutations in ITPR1 cause Gillespie syndrome with evidence for dominant-negative effect. Am J Hum Genet 2016;98:981–992.
33. Harms FL, Girisha KM, Hardigan AA, et al. Mutations in EBF3 disturb transcriptional profiles and cause intellectual disability, ataxia, and facial dysmorphism. Am J Hum Genet 2017;100:117–127.
34. Garcia-Gonzalo FR, Corbit KC, Sirerol-Piquer MS, et al. A transition zone complex regulates mammalian ciliogenesis and ciliary membrane composition. Nat Genet 2011;43:776–784.
35. Savukoski M, Klockars T, Holmberg V, Santavuori P, Lander ES, Peltonen L. CLN5, a novel gene encoding a putative transmembrane protein mutated in Finnish variant late infantile neuronal ceroid lipofuscinosis. Nat Genet 1998;19:286–288.
36. Krude H, Schutz B, Biebermann H, et al. Choreoathetosis, hypothyroidism, and pulmonary alterations due to human NKX2-1 haploinsufficiency. J Clin Invest 2002;109:475–480.
37. Tonelli A, D'Angelo MG, Salati R, et al. Early onset, non fluctuating spinocerebellar ataxia and a novel missense mutation in CACNA1A gene. J Neurol Sci 2006;241:13–17.
38. Brashear A, Mink JW, Hill DF, et al. ATP1A3 mutations in infants: a new rapid-onset dystonia-Parkinsonism phenotype characterized by motor delay and ataxia. Dev Med Child Neurol 2012;54:1065–1067.
39. Dard R, Mignot C, Durr A, et al. Relapsing encephalopathy with cerebellar ataxia related to an ATP1A3 mutation. Dev Med Child Neurol 2015;57:1183–1186.

Data available from Dryad (Additional References, References e1 to e5): links.lww.com/NXG/A270.

Genetic background of ataxia in children younger than 5 years in Finland

Erika Ignatius, Pirjo Isohanni, Max Pohjanpelto, et al.

Neurol Genet 2020;6;
DOI 10.1212/NXG.0000000000000444