Article

Heterozygous *UCHL1* loss-of-function variants cause a neurodegenerative disorder with spasticity, ataxia, neuropathy, and optic atrophy

Joohyun Park,[1,36] Arianna Tucci,[2,36] Valentina Cipriani,[2,3,4,5,36] German Demidov,[1] Clarissa Rocca[6], Jan Senderek,[7] Michaela Butryn,[8] Ana Velic,[9] Tanya Lam,[10,11] Evangelia Galanaki,[12] Elisa Cali,[12] Letizia Vestito,[2] Reza Maroofian,[12] Natalie Deininger,[1] Maren Rautenberg,[1] Jakob Admard,[1] Gesa-Astrid Hahn,[13] Claudius Bartels,[14] Nienke van Os,[15] Rita Horvath,[16] Patrick F. Chinnery,[16, 17] May Yung Tiet,[16] Channa Hewamadduma,[18,19] Marios Hadjivassiliou,[19, 20] George K. Tofaris,[21] Genomics England Research Consortium, Nicholas W. Wood,[22] Stefanie Hayer,[23,24] Friedemann Bender,[23,24] Benita Menden,[1] Isabell Cordts,[1,25] Katrin Klein,[1] Huu Phuc Nguyen,[26] Joachim K. Krauss,[27] Christian Blahak,[28,29] Tim M. Strom,[30,31] Marc Sturm,[1] Bart van de Warrenburg,[15] Holger Lerche,[32] Boris Maček,[9] Matthis Synofzik,[23,24] Stephan Ossowski,[1] Dagmar Timmann,[33] Marc E. Wolf,[29,34] Damian Smedley,[2] Olaf Riess,[1,35] Ludger Schöls,[23,24,35*] Henry Houlden,[2,37**] Tobias B. Haack,[1,35,37] Holger Hengel,[1,23,24,37]

[1]Institute of Medical Genetics and Applied Genomics, University of Tübingen, 72076 Tübingen, Germany

[2]William Harvey Research Institute, School of Medicine and Dentistry, Queen Mary University of London, London EC1M 6BQ, UK

[3]UCL Institute of Ophthalmology, University College London, London EC1V 9EL, UK

[4]Moorfields Eye Hospital NHS Foundation Trust, London EC1V 2PD, UK

[5]UCL Genetics Institute, University College London, London WC1E 6AA, UK

[6]Department of Neuromuscular Diseases, University College London Queen Square Institute of Neurology, London WC1N 3BG, UK

[7]Friedrich-Baur-Institute at the Department of Neurology, University Hospital, LMU Munich, 80336 Munich, Germany

[8]German Center for Neurodegenerative Diseases (DZNE), 39120 Magdeburg, Germany

[9]Proteome Center Tübingen, University of Tübingen, 72076 Tübingen, Germany

[10]Great Ormond Street Hospital NHS Trust, London WC1N 3JH, UK

[11]St George's Hospital NHS Trust, London SW17 0QT, UK

[12]Department of Neuromuscular Diseases, University College London Queen Square Institute of Neurology, London WC1N 3BG, UK

[13]CeGaT GmbH, Center for Genomics and Transcriptomics, 72076 Tübingen, Germany

[14]Department of Neurology, Otto-von-Guericke University, 39120 Magdeburg, Germany

[15]Department of Neurology, Donders Institute for Brain, Cognition and Behavior, Center of Expertise for Parkinson and Movement Disorders, Radboud University Medical Center, 6525 GA Nijmegen, The Netherlands

[16]Department of Clinical Neurosciences, University of Cambridge, John Van Geest Cambridge Centre for Brain Repair, Cambridge Biomedical Campus, Cambridge CB2 0PY, UK

[17]MRC Mitochondrial Biology Unit & Department of Clinical Neurosciences, University of Cambridge, Cambridge Biomedical Campus, Cambridge CB2 0QQ, UK

[18]Sheffield Institute for translational neurosciences (SITRAN), University of Sheffield, Sheffield S10 2HQ UK

[19]Royal Hallamshire Hospital, Sheffield Teaching Hospitals Foundation Trust , Sheffield S10 2JF, UK

[20]Academic Department of Neurosciences, Sheffield Teaching Hospitals NHS Trust and University of Sheffield, Sheffield S10 2JF, UK

[21]Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford OX3 9DU, UK

[22]Department of Clinical and Movement Neurosciences, University College London Queen Square Institute of Neurology, London, UK.

[23]Department of Neurodegenerative Diseases, Center for Neurology and Hertie-Institute for Clinical Brain Research, University of Tübingen, 72076 Tübingen, Germany

[24]German Center of Neurodegenerative Diseases (DZNE), 72076 Tübingen, Germany

[25]Department of Neurology, Klinikum rechts der Isar, Technical University Munich, 81675 Munich, Germany

[26]Department of Human Genetics, Medical Faculty, Ruhr University Bochum, 44801 Bochum, Germany

[27]Department of Neurosurgery, Hannover Medical School, 30625 Hannover, Germany

[28]Department of Neurology, Ortenau Klinikum Lahr-Ettenheim, 77933 Lahr, Germany

[29]Department of Neurology, Universitätsmedizin Mannheim, Medical Faculty Mannheim, University of Heidelberg, 68167 Mannheim, Germany

[30]Institute of Human Genetics, Technische Universität München, 81675 Munich, Germany

[31]Institute of Human Genetics, Helmholtz Zentrum München, 85764 Neuherberg, Germany

[32]Department of Neurology and Epileptology, Hertie Institute for Clinical Brain Research, University of Tübingen, 72076 Tübingen, Germany

[33]Department of Neurology and Center for Translational Neuro- and Behavioral Sciences (C-TNBS), Essen University Hospital, University of Duisburg-Essen, 45147 Essen, Germany

[34]Department of Neurology, Klinikum Stuttgart, 70174 Stuttgart, Germany

[35]Center for Rare Diseases, University of Tübingen, 72076 Tübingen, Germany

[36]These authors contributed equally to this work

[37]These authors contributed equally to this work

Correspondence to:

*Ludger Schöls, MD, Department of Neurodegenerative Diseases, Center for Neurology and Hertie-Institute for Clinical Brain Research, University of Tübingen, Otfried-Müller-Straße 27, 72076 Tübingen, Germany. E-mail: ludger.schoels@uni-tuebingen.de

**Henry Houlden, MD, PhD, Department of Neuromuscular Disorders, Institute of Neurology, University College London, Queen Square, WC1N 3BG London, UK. E-mail: h.houlden@ucl.ac.uk

Genetics in Medicine:

Abstract: 184 words (max 200)

Word count: 3998 words (max 4000)

Display items: 3 (max 5) and supplemental data

References: 40 (max 40)

ABSTRACT

Purpose: Bi-allelic variants in *UCHL1* (*Ubiquitin C-terminal hydrolase L1*) have been associated with a progressive early-onset neurodegenerative disorder, autosomal recessive spastic paraplegia type 79. Here, we investigate heterozygous *UCHL1* variants based on results from cohort-based burden analyses.

Methods: Gene burden analyses were performed on exome and genome data in independent cohorts of hereditary ataxia and spastic paraplegia patients from Germany and the UK on a total of 3,169 patients and 33,141 controls. Clinical data of affected individuals and additional independent families were collected and evaluated. Patients' fibroblasts were used to perform Mass-spectrometry-based proteomics.

Results: *UCHL1* was prioritized in both independent cohorts as a candidate gene for an autosomal dominant disorder. We identified a total of 34 cases from 18 unrelated families, carrying 13 heterozygous loss-of-function variants (15 families) and an inframe insertion (3 families). Affected individuals mainly presented with spasticity (24/31), ataxia (28/31), neuropathy (11/21) and optic atrophy (9/17). The mass spectrometry-based proteomics showed an approximately 50% reduction of UCHL1 expression in patients' fibroblasts.

Conclusion: Our bioinformatic analysis, in-depth clinical and genetic workup, and functional studies establish haploinsufficiency of UCHL1 as a novel disease mechanism in spastic ataxia.

Keywords: *UCHL1*, spastic ataxia, proteomics, gene burden

INTRODUCTION

Hereditary spastic paraplegias (HSPs) and hereditary ataxias (HAs) are rare neurogenetic diseases, that often show overlapping clinical features[1-4]. Besides the clinical overlaps, advances in next-generation sequencing have shown that there is also a shared genetic basis and molecular pathophysiology of these disorders. At least 69 genes with autosomal dominant, autosomal recessive, or X-linked inheritance have been described to cause disorders of a phenotypic spectrum overlapping HSPs and HAs[4].

Application of exome (ES) and genome sequencing (GS) in routine diagnostic setting drastically increased the diagnostic yield and the extent of generated genetic data. Yet it is estimated that ES leads to a genetic diagnosis only in half of the affected individuals with HSPs and/or HAs.[5,6] Analytic strategies in routine diagnostic settings are mostly single family/individual-based and do not usually involve large cohorts for the identification of rare pathogenic genetic variation. Data sharing and increasing availability of large rare disease cohorts open the perspective to realize the potential of cohort-based burden analyses even for these rare disorders. Here, we used case-control gene-burden analyses in independent cohorts of patients with HSPs and HAs, and discovered *ubiquitin C-terminal hydrolase L1* (*UCHL1*) as a gene for an autosomal dominant neurodegenerative disorder.

UCHL1 has previously been associated with autosomal recessive spastic paraplegia type 79 (SPG79), characterized by early-onset cerebellar ataxia, spastic paraplegia, and optic atrophy.[7-9] It encodes a deubiquitinating enzyme, that is neuron-specific and one of the most abundantly expressed proteins in the brain, constituting approximately 1 to 2% of soluble protein.[10] It is essential for maintaining ubiquitin homeostasis and its absence leads to aggregation of ubiquitinated proteins which is a common hallmark for many neurodegenerative diseases.[11,12] The role of *UCHL1* variants in numerous neurodegenerative disorders such as Alzheimer's disease (AD), Parkinson's disease (PD) and amyotrophic lateral sclerosis (ALS), have been widely explored, but the exact impact of *UCHL1* on the pathogenesis of these diseases is unclear.[13,14]

We identified 33 participants carrying 12 heterozygous predicted loss-of-function variants (LoFs) in 14 families and a highly predicted to be pathogenic inframe insertion in 3 families in *UCHL1*, and establish *UCHL1* to be associated with an autosomal dominant neurodegenerative disorder characterized by late-onset spastic-ataxia, neuropathy, and often optic atrophy.

MATERIALS, SUBJECTS AND METHODS

Genetic investigation

German cohort: diagnostic ES and GS were conducted in routine clinical care at the Institute of Medical Genetics and Applied Genomics Tübingen (Tübingen, Germany) as described previously.[15,16] Exome and genome sequencing libraries were generated from genomic DNA using Agilent SureSelect XT Human All Exon V5/V7 enrichment kits (Agilent Technologies, Santa Clara, CA, USA) or the TruSeq DNA PCR-Free k5it (Illumina, San Diego, CA, USA), respectively. Sequencing was performed on a HiSeq2500 [2 x 125, 2 x 100 base pairs (exome); Illumina, San Diego, CA, USA] or NovaSeq 6000 system [2 x 150 base pairs (genome); Illumina, San Diego, CA, USA] as paired-end reads. The sequence data was analyzed using the megSAP pipeline (https://github.com/imgag/megSAP) and aligned to the GRCh37 reference genome. Prior to burden analysis all datasets have been analyzed in a diagnostic setting according to an in-house standard operating procedure including a variety of different filtering steps to prioritize likely clinically relevant DNA variants. Identified variants were classified following the recommendations of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology.[17]

U.K. cohort within the 100KGP: DNA extracted from blood was prepared for GS using TruSeq DNA PCR-Free library preparation. 150 or 125 bp paired-end sequencing was performed on either HiSeq 2000 or HiSeq X platforms. Genomes were sequenced to an average minimum depth of 35X (31X - 37X); sequencing reads were aligned to the human genome reference build 37 or build 38 using of Isaac Genome Alignment Software. Family-

based variant calling of single-nucleotide variants (SNVs) and insertion or deletions (indels) was performed with the use of the Platypus variant caller.

Gene burden analysis

German cohort: case-control gene burden analysis was first performed on 1,547 selected diagnostic cases from the in-house database at the Institute of Medical Genetics and Applied Genomics Tübingen presenting with either spastic paraplegia, ataxia, or spastic ataxia and 3,624 controls (i.e. healthy controls and patients with non-related diseases without HSP or HA) using burden test implemented in the rvGWAS package.[18] Related family members were excluded from both groups. Genetic ancestry was inferred using the SampleAncestry software (https://github.com/imgag/ngs-bits). Due to limited availability of different ancestries among controls, only European descent samples were used for this analysis. This preselection might represent a putative limitation regarding diversity and generalizability of our approach. Male and female ratio was matched between the case and control cohorts. Non-synonymous variants, including splice site alterations (±20 nucleotides from exon-intron boundary) involving any of the ~ 20,000 consensus coding sequence (CCDS) genes, were extracted and subsequently filtered using the following by filtering criteria. Only variants with at least 20x coverage, a minor allele frequency (MAF) < 0.5% in gnomAD (v2.1.1 and v3.1.1, Genome Aggregation Database, gnomad.broadinstitute.org) and < 2.5% in the in-house database, with high impact (annotated as start-lost, stop-gain, frameshift, and canonical splice site alterations located ± 2 nucleotides) or high pathogenicity scores (CADD score > 25 or SIFT/PolyPhen damaging annotation) were included in the analysis. To control the false discovery rate (FDR), the overall *p* value from the gene burden testing was adjusted using the Benjamini-Hochberg procedure. Subsequently, enrichment of *UCHL1* variants in HSPs and HAs was evaluated in the full in-house dataset of 14,303 exomes and genomes at the Institute of Medical Genetics and Applied Genomics Tübingen. At the time of the study, 2,066 out of 14,303 individuals were reported to show spasticity and/or ataxia. Among them, 272 were described to present both spastic paraplegia and ataxia. The enrichment of rare variants in cases was analyzed by one-

sided Fisher's exact test using R statistical software (https://www.r-project.org/). A *p* value $< 2.5 \times 10^{-6}$ was used to claim gene-based genome/exome-wide significance corrected for ~20,000 CCDS genes.

U.K. cohort within the 100KGP: case-control gene burden analysis was performed within the rare disease component of the 100KGP.[19] Cases were defined as all 100KGP probands recruited under a certain clinical indication, e.g. 'hereditary ataxia', n = 1,103, while corresponding 'controls' were all remaining recruited 100KGP probands except those recruited under the broader disease category of the clinical indication, e.g. 'neurological and neurodevelopmental disorder', n = 20,904. The analysis was limited to those disease gene associations where at least five cases exist for the clinical indication and where relevant variants in the gene were seen in at least four probands over the entire cohort of cases and controls. Exomiser was then run on all probands' genome data to filter (i) rare (MAF < 0.1%, for dominant, and <1%, for recessive, in gnomAD v2.1.1 and v3.1.1 as well as within the local 100KGP cohort)), (ii) coding variants, (iii) that segregated with the disease status (if multiple family members were available) as expected for each possible mode of inheritance.[20] Gene-based enrichment of rare variants in cases was assessed using one-sided Fisher's exact test (as implemented in R statistical software) under four scenarios: (1) enrichment of rare, predicted LoFs, (2) enrichment of rare, predicted pathogenic variants (with Exomiser variant score > 0.8, that is variants that are predicted to be pathogenic by in silico prediction tools REVEL and/or MVP), (3) enrichment of rare variants with Exomiser variant score > 0.8 and in a constrained coding region, (4) enrichment of rare, *de novo* variants.[21] For the latter, only trios or larger families where *de novo* calling was possible were considered. The Benjamini-Hochberg procedure was used to correct for multiple testing; an overall FDR adjusted q value threshold of 0.10 was used for claiming significant gene-disease associations considering the total number of tests under all four scenarios and all 197 different 100KGP clinical indications analysed (621,741 tests).

All 100KGP genomes with HA or HSP had been screened for copy number variants (CNVs) and short tandem repeats (STRs) in HA and HSP genes from the PanelApp virtual gene panels.[22] CNV calls were generated by Genomics England using Manta and Canvas software (Illumina); STR calls were generated using ExpansionHunter.[22,23] No likely pathogenic variants were identified.

Clinical investigation

Patients from Germany were examined and phenotyped by experienced neurologists and referred to the genetic center at the University of Tübingen for diagnostic ES or GS. Additional patients were recruited through communication with collaborators and the RD-Connect Genome-Phenome Platform (https://platform.rd-connect.eu/). Detailed clinical data of individuals with LoFs in *UCHL1* were obtained and analyzed.

Patients from the U.K. with genetically undiagnosed neurological disease were recruited to the 100KGP by neurologists at English hospitals. At recruitment, standardized clinical data were recorded using the Human Phenotyping Ontology, according to disease-specific data model (https://www.genomicsengland.co.uk/?wpdmdl=5500).[24] Following the identification of families carrying *UCHL1* LoFs with spastic ataxia by gene burden analysis, additional information was collected retrospectively by contacting the recruiting clinician for each patient (Table S5).

Fibroblast cultivation and protein isolation

Human dermal fibroblasts were maintained at 37°C, 5% CO_2 and 100% relative humidity in fibroblast medium consisting of Dulbecco's Modified Eagle Medium (Merck) supplemented with 10% foetal bovine serum (Thermo Fisher Scientific) in cell culture flasks. Upon reaching high confluence, primary fibroblasts were washed with cold PBS, scraped off in PBS, centrifuged at 800 g for 5 min and frozen at -80°C. Pellets of primary fibroblasts were lysed in RIPA buffer (Sigma) containing protease inhibitors (Roche) for 45 min on a rotator at 4°C. Cell debris were pelleted at 15,800 g and 4°C for 30 min. The protein concentration was

determined using the Pierce BCA Protein Assay Kit (Thermo Fisher Scientific) according to the manufacturer's instructions.

NanoLC-MS/MS analysis

Protein samples for nanoLC-MS/MS were generated from five fibroblast cell lines in biological duplicates (F1-IV.1, F1-IV.3, F2-III.2, F2-IV.1, F3-II.1) and compared to five control fibroblast cell lines (age and gender-matched) in biological duplicates (Table S7). Nano LC MS/MS analysis and data processing were performed as previously described.[25]

RESULTS

Gene burden analysis identifies an excess of heterozygous LoFs in families with spasticity or ataxia.

To detect enrichment for predicted pathogenic variants in novel genes for ataxia and spastic paraplegia, gene burden analyses were performed independently within a German cohort by a research group in Tuebingen (Germany) and within the 100KGP by researchers in London (U.K.), on a total of 3,169 patients and 33,141 controls across all cohorts.

For the German cohort, gene-based collapsing analysis (in 1,547 cases and 3,624 controls) indicated an enrichment of rare variants in *SPAST*, *KIF5A*, *UCHL1*, *POLR3A* and *FIGN* (Table S1). While *SPAST* and *KIF5A,* which were top-listed, are dominant-acting genes, *UCHL1* and *POLR3A,* have been associated with autosomal recessive disorders.[26,27] To our knowledge, variants in *FIGN* have so far not been associated with human phenotypes. Follow-up of the identified in-house individuals carrying rare *FIGN* variants did support the hypothesis a specific disease association. Remarkably, the variants that generated the *UCHL1* signal were LoFs and did not derive from possible recessive *UCHL1* cases. A subsequent *UCHL1* enrichment analysis in the full in-house database containing 14,303 datasets (2,066 cases with HSP and/or HA, and 12,237 controls) revealed a significant enrichment of LoFs in cases reaching exome-wide significance when compared to gnomAD controls (Table S2). The

enrichment was more significant among individuals with both HSP and HA (p = 6.536 x 10^{-8}, 3/272 cases vs. 3/182,146 gnomAD controls) than in the overall group with individuals with HSP and/or HA (p = 5.375 x 10^{-7}, 4/2,066 cases vs. 3/182,146 gnomAD controls). Notably, the observed/expected (o/e) number of LoFs in gnomAD was 0, indicating a high intolerance of *UCHL1* for haploinsufficiency. Rare missense variants in *UCHL1* (MAF < 0.1%) from our total in-house database were not enriched in cases (Table S3).

UCHL1 had independently been investigated as candidate gene for the clinical indication HA within the follow-up of the cohort-wise gene burden analysis in the rare disease component of the 100KGP (57,002 genomes sequenced from 27,591 families affected by rare diseases).[19] A new gene burden analysis within the latest available 100KGP dataset at the time of analysis (March 2021) of 35,422 rare disease families, including 1,103 HA probands used as cases and 20,904 non-neurological probands used as controls, revealed a statistically significant burden of either LoFs (p = 1.5 x 10^{-5}, q = 0.0231, 5/1,103 cases vs. 3/20,904 controls, Table S2) or any variants highly predicted to be pathogenic (including LoFs), i.e. Exomiser variant score >= 0.80, (p = 2.0 x 10^{-5}, q = 0.0265, 6/1,103 cases vs. 7/20,904 controls) in *UCHL1*. All heterozygous LoFs identified in cases were absent from control cohorts and gnomAD, and none of them have been previously reported in an autosomal recessive context.

Fifteen families carry 13 LoFs and 3 families the same inframe insertion in *UCHL1*

Figures 1, 2 and Table S5 summarize the *UCHL1* variants and clinical phenotype. Within the German cohort, four different heterozygous LoFs from four independent, so far genetically unsolved families (F1-4, Figure 1a) were identified through gene burden analysis. In family F1, the index F1-III.3 carried a frameshift variant c.64dup, p.(Val22glyfs*39) in *UCHL1.* Subsequently, four affected and one healthy family member were recruited (Figure 1a). F1-IV.3 had been submitted independently for diagnostic ES, and by extending the family history, we were able to reconstruct this individual's relationship to F1. In F2, the frameshift variant, c.349_364del, p.(Phe117Argfs*33), identified in F2-III.2, and segregated in three healthy

sisters (F2-III.1, F2-III.6 and F2-III.7) and two affected brothers (F2-III.5 and F2-III.8). One affected brother (F2-III.5) had an additional diagnosis of ALS and carried a heterozygous pathogenic *SOD1* variant, c.272A>C, p.(Asp91Ala). The daughter (F2-IV.1) of a similarly affected sister (F2-III.3) was subsequently diagnosed with ataxia and the same variant was identified independently by diagnostic GS. In F3, a heterozygous c.385_388dup, p.(Ala130Glufs*6) variant was identified *de novo* in F3-II.1, and the healthy sister did not carry the variant. Finally, a splice variant c.527-1G>A, p.(?) was identified in a singleton in F4 (F4-II.1).

Furthermore, an additional family (F5) carrying a c.387_388delAG, p.(Arg129Serfs*5) was identified through the RD-Connect Platform and two other families (F6 and F7) by communication with collaborators. Two affected sisters and the affected father from F6 (F6-III.1, F6-III.3 and F6-II.1) segregated for the LoF variant c.73C>T, p.(Gln25*). The same variant was found in a different unrelated neurological patient (F7-II.1) via diagnostic GS (Table S5). During the period of revision, we identified a different nonsense variant, c.4C>T, p.(Gln2*) in a 54 year old patient (F18) with HSP and ataxia by diagnostic GS.

In the U.K. cohort, 10 autosomal dominant families carrying either a LoF variant (n = 7) or a variant highly predicted to be pathogenic (n = 3) were altogether identified for *UCHL1*. First, five HA families with LoFs were identified, including F8, a male participant with an autosomal dominant family history of ataxia carrying the c.381-384del, p.(Asp128Glufs*26) variant; F9, a duo with affected father and son both carrying the c.532C>T, p.(Arg178*) variant; F10, a singleton carrying the c.631G>T, p.(Glu211*) variant; F11, a singleton carrying the c.260_261dup, p.(Asn88Glyfs*28); F12, a duo with affected mother and daughter both carrying the c.95_98dup, p.(Leu34Glyfs*28) variant, and an autosomal dominant family history of disorder. The phenotypic and family data recorded in these families was consistent with autosomal dominant late-onset spastic ataxia with visual involvement. This prompted the search and identification of additional families in the 100KGP cohort, as detailed below.

F13, a duo family with father and son with a dominant family history, recruited as HSP, were identified as carrying the same variant found in family F9, c.532C>T, p.(Arg178*). In F14, a proband recruited under inherited optic neuropathy was found to carry the c.116_117del, p.(Leu39Argfs*21), also present in the mother, initially recorded as unaffected. The clinical re-examination of the family revealed an autosomal dominant family history of ataxia, with the mother and the maternal grandfather affected. Three more families (F15, F16, F17) were identified when assessing an excess in cases of variants highly predicted to be pathogenic, i.e. Exomiser variant score >= 0.80. That first led to the identification of a proband with HA in family F15 carrying an inframe duplication variant c.154_156dup, p.(Leu52dup); the same variant was also found in the proband from family 16 (F16-II.1) recruited under HSP, and in another proband, his sister and his daughter (F17-II.I, F17-II.3, F17-III.1) recruited under HA and inherited optic neuropathy (Table S5). The c.154_156dup, p.(Leu52dup) variant present in those three families is absent in gnomAD and in the 100KGP dataset.

Clinical phenotypes of *UCHL1* families involve spasticity, ataxia, neuropathy and/or optic atrophy

Detailed clinical information was available in 31 of 33 individuals from 18 families carrying 13 LoFs and one inframe duplication, highly predicted to be pathogenic, in *UCHL1*. On clinical examination (Figure 2, Table S5), lower limb spasticity with brisk deep tendon reflexes was observed in 24 of 31 individuals. Gait ataxia was present in 28 individuals and 26 of them had predominantly cerebellar ataxia with either saccadic eye movements (16/31), dysarthria or dysphagia (9/31), and/or intention tremor (18/31). Tremor other than intention tremor was reported in 6 of 31 individuals. Two individuals presented with predominantly sensory ataxia. Reduced surface sensibility or pallesthesia in the lower limbs was observed in 24 individuals, and sensory or sensorimotor axonal neuropathy could be demonstrated electrophysiologically in 11 of 21 examined patients. Ten patients reported visual impairment, and optic atrophy was detected in 9 of 17 examined individuals. Disease onset varied from childhood to age 70 (median 49) with slow progression of the movement disorder according to

medical history. In summary, the core phenotype consists of cerebellar ataxia (26/31), spasticity (24/31), sensory or sensorimotor neuropathy (11/21) and optic atrophy (9/17) (Figure 2, Table S5).

Some individuals differed from this core phenotype. F1-III.4 suffered from generalized dystonia from the age of 53 years onwards and developed, possibly independent from dystonia, a spastic-ataxic gait disorder at the age of 70. F2-III.5 presented lower motoneuron signs in addition to a spastic-ataxic phenotype, clinically compatible with a slowly progressive ALS possibly caused by the additional heterozygous *SOD1* variant (c.272A>C, p.[Asp91Ala]).[28] The most severely affected individual was F3-II.1. He had a history of childhood-onset developmental delay, followed by optic atrophy with severe visual impairment (visus c.c. 0.06 on both eyes) diagnosed around the age of 12 years. Spastic-ataxic gait disorder developed at the age of 25 years.

One individual (F12-III.2) also presented with early-onset visual difficulties (at 7 years) and was diagnosed with optic atrophy in his twenties. He is currently 24 years old and does not present any signs of spasticity or ataxia, unlike his mother (F12-II.1).

Exploring dysregulated genes in proband-derived fibroblasts

In the nanoLC-MS/MS data, 75 proteins were significantly up- or down-regulated (log Student's T-test p-value ≤ 0.05 and absolute |log FC ≥ 0.5|, Table S6) and the protein levels of UCHL1 were significantly decreased (-2.4 logFC, Figure 3a). Remarkably, this unbiased approach showed UCHL1 as the protein with the largest decrease in protein level (Figure 3a). An interaction network based on the 30 most up- or downregulated proteins was computed using string-db (https://string-db.org/), revealing a significant enrichment of interactions among these proteins (protein-protein interaction enrichment p = 1.17e-7, Figure 3b). Of note, the protein with the highest increase in protein level (2.05 logFC) was MME (membrane metallo-endopeptidase or neprilysin) which was directly linked to UCHL1 (Figure 3b).

DISCUSSION

In this study, we provide several lines of evidence that heterozygous LoFs in *UCHL1* are the genetic cause of an autosomal dominant inherited neurodegenerative disorder with a distinct phenotype. Two independent case-control gene burden analyses prioritized *UCHL1* among the top hits. Surprisingly the signal was not generated by previously described missense variants that were associated with autosomal recessive SPG79[7-9], but by heterozygous LoFs.

In total, 13 different heterozygous *UCHL1* LoFs were identified in 29 patients from 15 families showing autosomal dominant inheritance and a common core phenotype. This core phenotype consists of a movement disorder with spasticity, cerebellar and/or sensory ataxia, and often sensorimotor neuropathy. Visual impairment and optic atrophy were other frequent features. While some participants had early-onset visual impairment, others only had subclinical optic atrophy or no optic atrophy at all. During extended screening for *UCHL1* variants, we also identified three independent families with the heterozygous inframe duplication c.154_156dup, p.(Leu52dup). Participants from all three families (F15-II.I, F16-II.I, F17-II.I, F17-II.3 and F17-III.I) presented with the specific combination of optic atrophy and late-onset spastic-ataxia. Therefore, we consider it likely that this variant is also pathogenic.

Previously, bi-allelic variants in *UCHL1* have been described to cause SPG79 in ten individuals from four families. A consanguineous family of Turkish origin with three affected siblings harboring a homozygous missense variant (c.20A>C, p.[Glu7Ala]) presented an early-onset progressive neurodegenerative syndrome with visual loss, ataxia, spasticity, and cognitive impairment.[7] A Norwegian family with compound heterozygous missense variants (c.533G>A, p.[Arg178Gln] and c.647C>A, p.[Ala216Asp]) and an Indian family with the homozygous splice variant (c.459+2T>C, p.[?]) demonstrated a similar phenotype.[8,9] More recently, a homozygous inframe deletion c.627_629del, p.(Gly210del) has been reported in two siblings, showing childhood-onset visual impairment, progressive spasticity, and ataxia, compatible with the diagnosis of Behr syndrome.[29] SPG79 patients had an early onset of disease between 2 and 10 years (median 7.5), and a more progressive disease course leading to wheel-chair dependency and often blindness. Three patients became non-ambulatory in

early childhood between 5 and 9 years, while some others in late adulthood after three to five decades.[7-9,29] The deceased mother of the Norwegian family reportedly had late onset tremor. However, one of the wild-type siblings also showed tremor and one sister carrying a heterozygous p.Arg178Gln variant did not show any neurological signs at the age of 74 years. Therefore, it remains unclear whether the tremor symptom is related to a heterozygous *UCHL1* variant. Information on the Indian patients' parents was not available.[8] All other parents were reported as asymptomatic and their affected children were last examined between 28 and 65 years of age. [7,9,29] In summary, identical neurological systems are affected in both the previously reported recessive SPG79 and the currently reported cohort with heterozygous *UCHL1* variants (Figure 2), while the currently reported probands have a milder course with a later disease onset ranging from 12 to 70 years of age (median 49) on average.

The functional data initially suggested loss-of-function as pathomechanism for SPG79. UCHL1 knockout mice displayed spastic gait movements and progressive paralyses in their hind limbs [30], and the missense variant p.(Glu7Ala) showed reduced enzymatic hydrolase activity of <10%.[7] However, the two compound heterozygous missense variants detected in SPG79 patients had two opposing functional consequences. While p.(Ala216Asp) proteins were degraded, the p.(Arg178Gln) had enhanced enzymatic activity. The authors proposed that the increased enzymatic activity on one allele may have a protective influence on cognitive dysfunction.[9] Finally, the splice variant (c.459+2T>A, p.[?]) is predicted to cause a skipping of exon 6, which would create an inframe deletion of 16 amino acids.[8,29] In total, only missense variants, inframe-deletions and one splice variant predicted to cause an inframe deletion have been associated with SPG79. Since the parents of SPG79 patients were reportedly asymptomatic, it is possible that the mutant UCHL1 gene product still has relevant residual function and that the disease mechanism differs from loss-of-function / haploinsufficiency.

Our mass spectrometry-based proteomics data showed significantly decreased UCHL1 levels by approximately 50%, indicating haploinsufficiency of *UCHL1* as the most likely pathomechanism for the autosomal dominant disorder. Regarding other dysregulated proteins,

our pathway analysis pattern revealed a direct link between UCHL1 and the most upregulated protein, MME. Pathogenic variants in *MME* have been associated with Charcot-Marie-Tooth disease type 2T and autosomal dominant spinocerebellar ataxia type43.[31,32] MME cleaves and inactivates several peptides and is known to degrade amyloid-beta (Aß). UCHL1 has been well described to inactivate the Beta-Secretase 1 (BACE1), the major beta-secretase for the generation of Aβ peptides from APP.[33] Since both MME and UCHL1 are connected with the metabolism of amyloid-β precursor protein (APP) and the clearance of Aβ, we speculate that the overexpression of MME could be in part a compensatory mechanism. By including APP in addition to the 30 most up- or down-regulated proteins in the network analysis, it appears that APP acts as a central connecting node linking several of the dysregulated proteins in the interaction network (Figure S1).

UCHL1 has previously been linked to various other more common neurodegenerative disorders. In detail, two missense variants in *UCHL1* have been implicated in PD. While the variant p.(Ile93Met) was described to cause early-onset PD with incomplete penetrance, the polymorphism p.(Ser18Tyr) appeared to be a protective factor for PD.[14,34] However, both theories have been negated by several subsequent meta-analyses.[35] Some data indicated that UCHL1 might be involved in the pathophysiology and disease progression of AD and ALS. Reduced UCHL1 expression was observed in individuals with AD, and overexpression of UCHL1 improved beta-amyloid-induced synaptic dysfunction and memory deficits in AD mouse models.[13,36,37] Furthermore, quantitative proteomic studies revealed elevated UCHL1 levels in cerebrospinal fluid of individuals with ALS, suggesting UCHL1 as a potential biomarker that might be involved in the disease progression of ALS.[38]

Overall, the exact pathomechanism of *UCHL1* variants in recessive disorders and in other neurodegenerative diseases such as PD and AD remains unclear. Notably, natural and synthetic antisense to mouse *Uchl1* (AS-*Uchl1*) have shown to activate *UCHL1* mRNA translation increasing protein synthesis at post-translational level *in vitro*.[39] On the basis of this discovery, the use of AS Uchl1-derived long non-coding RNAs, named SINEUPs, has been proposed as potential new RNA-based therapy for gene-specific haploinsufficiency diseases.[40]

In this regard, the use of AS-*Uchl1* could be a potential therapy for patients harboring LoFs in *UCHL1.*

The statistical analyses, supporting segregation data in large dominant pedigrees, and the proteomics data provide evidence of *UCHL1* haploinsufficiency as the pathomechanism underlying an autosomal dominant neurological disorder presenting with spasticity, cerebellar and/or sensory ataxia, polyneuropathy, and often optic atrophy. The disease presentation is similar but milder than the previously reported recessive disorder SPG79. Continuous exploration of *UCHL1* variants (either in a heterozygous or homozygous state) and its function may provide further insight into the pathogenesis of various neurodegenerative diseases.

SUPPLEMENTAL DATA

Figure S1: Interaction network based on the 30 most over- or underexpressed from nanoLC-MS/MS using string-db (https://string-db.org/). Additionally, APP was added to the network analysis and appeared to be a central connection node between the dysregulated proteins. Overexpressed proteins are shown in red, underexpressed proteins in blue.

Table S1: Gene-based rare-variant burden analysis results using 1,547 diagnostic cases with either spastic paraplegia, ataxia or both phenotypes and 3,624 controls in the German cohort.

Table S2: Counts of different *UCHL1* LoF alleles (transcript NM_004181.4/ENST00000284440.4) in the German and U.K. cohorts, and in the gnomAD datasets (v2.1.1 and v3.1.1, non-neuro).

Table S3: Counts of different individuals with *UCHL1* missense variants (transcript NM_004181.4/ENST00000284440.4, with MAF < 0.1% in gnomAD) in the German cohort, and in the gnomAD datasets (v2.1.1 and v3.1.1, non-neuro).

Table S4: Gene variant position and population frequency of *UCHL1* (NM_004181.4/ENST00000284440.4) LoFs identified in 14 families and an inframe insertion identified in 3 families.

Table S5: Clinical and Genetic Findings in Individuals with heterozygous *UCHL1* Variants

Table S6: Results from nanoLC-MS/MS based on fibroblasts from F1-IV.1, F1-IV.3, F2-III.2, F2-IV.1, F3-II.1 in comparison to controls

ETHICAL APPROVAL

German cohort: written informed consent was obtained from all individuals or their guardians. The study was approved by the ethics committee of the medical faculty by the local Institutional Review Board of the Medical Faculty of the University of Tübingen, Germany (vote 598/2011 and 066/2021BO2).

U.K. cohort within the 100,000 Genomes Project (100KGP): following ethical approval from the the national research ethics committee (14/EE/1112), consent was obtained from all patients recruited to the 100KGP.[19]

CONFLICT OF INTEREST

The authors declare no competing interests.

DATA AVAILABILITY

Depersonalized data and additional experimental data that this study is based on, can be provided upon request

ACKNOWLEDGEMENTS

We thank all the families for their participation. AT is an MRC Clinician Scientist investigator (MR/S006753/1). LS, MS, OR, BvdW and HH are members of the European Reference Network for Rare Neurological Diseases (ERN-RND). TBH was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation: 418081722, 433158657). HH was supported by the intramural fortüne program (#2554-0-0) and by the DFG under the project number HE 8803/1–1. This project was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) No 441409627, as part of the PROSPAX consortium under the frame of EJP RD, the European Joint Programme on Rare Diseases, under the EJP RD COFUND-EJP N° 825575 (to MS, BvdW and – as an associated partner – DT) and by grant 779257 “Solve-RD” from the Horizon 2020 research and innovation programme (to MS, OR and BvdW). RH is a Wellcome Trust Investigator (109915/Z/15/Z), who receives support from the Medical Research Council (UK) (MR/N025431/1 and MR/V009346/1), the European Research Council (309548), the Newton Fund (UK/Turkey, MR/N027302/1), the Addenbrookes Charitable Trust (G100142), the Evelyn Trust, the Stoneygate Trust, the Lily Foundation and an MRC strategic award to establish an International Centre for Genomic Medicine in Neuromuscular Diseases (ICGNMD) MR/S005021/1. This research was supported by the NIHR Cambridge Biomedical Research Centre (BRC-1215-20014). The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care. PFC is a Wellcome Trust Principal Research Fellow (212219/Z/18/Z), and a UK NIHR Senior Investigator, who receives support from the Medical Research Council Mitochondrial Biology Unit (MC_UU_00015/9), the Medical Research Council (MRC) International Centre for Genomic Medicine in Neuromuscular Disease (MR/S005021/1), the Leverhulme Trust (RPG-2018-408), an MRC research grant (MR/S035699/1), an Alzheimer's Society Project Grant (AS-PG-18b-022). This research was supported by the NIHR Cambridge Biomedical Research Centre (BRC-1215-20014). The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care. DT received funding from the German Reseach

Foundation, EU and Bernd Fink Foundation, unrelated to the present study. GKT is supported by an MRC Senior Clinical Fellowship (MR/V007068/1).

AUTHOR INFORMATION

Conceptualization: J.P., A.T., V.C., H.Hengel, T.B.H, H.Houlden; Data curation: J.P., A.T., V.C., H.Hengel; Data analysis: J.P., A.T., V.C., H.Hengel, T.B.H, T.L., A.V., E.G., L.V., N.D., M.R., J.A., B.M., I.C., S.O., T.M.S., M.S., D.S.; Patients recruitment and phenotypic characterization: J.P., A.T., H.Hengel, L.S., H.Houlden., J.S., M.B., E.C., R.M., G.H., C.B., N.v.O., R.H., P.F.C., M.Y.T., C.H., M.H., G.K.T., N.W.W., S.H., F.B., J.K.K., C.B., B.vd.W., H.L., M.S., D.T., M.E.W.; Methodology: J.P., A.T., V.C., H.Hengel, G.D., C.R., A.V., B.M.; Supervision: L.S., H.Hengel, T.B.H, H.Houlden; Visualization: J.P., H.Hengel., J.A.; Writing-original draft: J.P., A.T., V.C., H.Hengel; Writing-review & editing: J.P., A.T., V.C., H.Hengel, O.R., T.B.H, L.S., H.Houlden. All authors revised the manuscript for intellectual content.

CONSORTIA

Members of The Genomics England Research Consortium:

J. C. Ambrose[1], P. Arumugam[1], E. L. Baple[1], M. Bleda[1], F. Boardman-Pretty[1,2], J. M. Boissiere[1], C. R. Boustred[1], H. Brittain[1], M. J. Caulfield[1,2], G. C. Chan[1], C. E. H. Craig[1], L. C. Daugherty[1], A. de Burca[1], A. Devereau[1], G. Elgar[1,2], R. E. Foulger[1], T. Fowler[1], P. Furió-Tarí[1], J. M. Hackett[1], D. Halai[1], A. Hamblin[1], S. Henderson[1,2], J. E. Holman[1], T. J. P. Hubbard[1], K. Ibáñez[1, 2], R. Jackson[1], L. J. Jones[1,2], D. Kasperaviciute[1,2], M. Kayikci[1], L. Lahnstein[1], K. Lawson[1], S. E. A. Leigh[1], I. U. S. Leong[1], F. J. Lopez[1], F. Maleady-Crowe[1], J. Mason[1], E. M. McDonagh[1,2], L. Moutsianas[1,2], M. Mueller[1,2], N. Murugaesu[1], A. C. Need[1,2], C. A. Odhams[1], C. Patch[1,2], D. Perez-Gil[1], D. Polychronopoulos[1], J. Pullinger[1], T. Rahim[1], A. Rendon[1], P. Riesgo-Ferreiro[1], T. Rogers[1], M. Ryten[1], K. Savage[1], K. Sawant[1], R. H. Scott[1], A. Siddiq[1], A. Sieghart[1], D. Smedley[1,2], K. R. Smith[1,2], A. Sosinsky[1,2], W. Spooner[1], H. E. Stevens[1], A. Stuckey[1], R. Sultana[1], E. R. A. Thomas[1,2], S. R. Thompson[1], C. Tregidgo[1], A. Tucci[1,2], E. Walsh[1], S. A. Watters[1], M. J. Welland[1], E. Williams[1], K. Witkowska[1,2], S. M. Wood[1,2], M. Zarowiecki[1]

[1]Genomics England, London, UK
[2]William Harvey Research Institute, Queen Mary University of London, London, EC1M 6BQ, UK

WEB RESOURCES

The URLs for the data presented herein are as follows:

gnomAD server, https://gnomad.broadinstitute.org/

GTEx Portal, https://gtexportal.org

Online Mendelian Inheritance in Man (OMIM), https://omim.org/

megSAP pipeline, https://github.com/imgag/megSAP

string-db (https://string-db.org/)

RD-Connect Genome-Phenome Platform, https://platform.rd-connect.eu/

REFERENCES

1. Bourassa, C.V., Meijer, I.A., Merner, N.D., Grewal, K.K., Stefanelli, M.G., Hodgkinson, K., Ives, E.J., Pryse-Phillips, W., Jog, M., Boycott, K., et al. (2012). VAMP1 mutation causes dominant hereditary spastic ataxia in Newfoundland families. Am J Hum Genet *91*, 548-552. 10.1016/j.ajhg.2012.07.018.
2. de Bot, S.T., Willemsen, M.A., Vermeer, S., Kremer, H.P., and van de Warrenburg, B.P. (2012). Reviewing the genetic causes of spastic-ataxias. Neurology *79*, 1507-1514. 10.1212/WNL.0b013e31826d5fb0.
3. Meijer, I.A., Hand, C.K., Grewal, K.K., Stefanelli, M.G., Ives, E.J., and Rouleau, G.A. (2002). A locus for autosomal dominant hereditary spastic ataxia, SAX1, maps to chromosome 12p13. Am J Hum Genet *70*, 763-769. 10.1086/338933.
4. Synofzik, M., and Schule, R. (2017). Overcoming the divide between ataxias and spastic paraplegias: Shared phenotypes, genes, and pathways. Movement disorders : official journal of the Movement Disorder Society *32*, 332-345. 10.1002/mds.26944.
5. Ngo, K.J., Rexach, J.E., Lee, H., Petty, L.E., Perlman, S., Valera, J.M., Deignan, J.L., Mao, Y., Aker, M., Posey, J.E., et al. (2020). A diagnostic ceiling for exome sequencing in cerebellar ataxia and related neurological disorders. Hum Mutat *41*, 487-501. 10.1002/humu.23946.
6. Schule, R., Wiethoff, S., Martus, P., Karle, K.N., Otto, S., Klebe, S., Klimpe, S., Gallenmuller, C., Kurzwelly, D., Henkel, D., et al. (2016). Hereditary spastic paraplegia: Clinicogenetic lessons from 608 patients. Ann Neurol *79*, 646-658. 10.1002/ana.24611.
7. Bilguvar, K., Tyagi, N.K., Ozkara, C., Tuysuz, B., Bakircioglu, M., Choi, M., Delil, S., Caglayan, A.O., Baranoski, J.F., Erturk, O., et al. (2013). Recessive loss of function of the neuronal ubiquitin hydrolase UCHL1 leads to early-onset progressive neurodegeneration. Proc Natl Acad Sci U S A *110*, 3489-3494. 10.1073/pnas.1222732110.
8. Das Bhowmik, A., Patil, S.J., Deshpande, D.V., Bhat, V., and Dalal, A. (2018). Novel splice-site variant of UCHL1 in an Indian family with autosomal recessive spastic paraplegia-79. J Hum Genet *63*, 927-933. 10.1038/s10038-018-0463-6.
9. Rydning, S.L., Backe, P.H., Sousa, M.M.L., Iqbal, Z., Oye, A.M., Sheng, Y., Yang, M., Lin, X., Slupphaug, G., Nordenmark, T.H., et al. (2017). Novel UCHL1 mutations reveal new insights into ubiquitin processing. Hum Mol Genet *26*, 1031-1040. 10.1093/hmg/ddw391.
10. Lowe, J., McDermott, H., Landon, M., Mayer, R.J., and Wilkinson, K.D. (1990). Ubiquitin carboxyl-terminal hydrolase (PGP 9.5) is selectively present in ubiquitinated inclusion bodies characteristic of human neurodegenerative diseases. J Pathol *161*, 153-160. 10.1002/path.1711610210.
11. Komander, D., Clague, M.J., and Urbe, S. (2009). Breaking the chains: structure and function of the deubiquitinases. Nat Rev Mol Cell Biol *10*, 550-563. 10.1038/nrm2731.

12. Jara, J.H., Frank, D.D., and Ozdinler, P.H. (2013). Could dysregulation of UPS be a common underlying mechanism for cancer and neurodegeneration? Lessons from UCHL1. Cell Biochem Biophys *67*, 45-53. 10.1007/s12013-013-9631-7.
13. Choi, J., Levey, A.I., Weintraub, S.T., Rees, H.D., Gearing, M., Chin, L.S., and Li, L. (2004). Oxidative modifications and down-regulation of ubiquitin carboxyl-terminal hydrolase L1 associated with idiopathic Parkinson's and Alzheimer's diseases. J Biol Chem *279*, 13256-13264. 10.1074/jbc.M314124200.
14. Nishikawa, K., Li, H., Kawamura, R., Osaka, H., Wang, Y.L., Hara, Y., Hirokawa, T., Manago, Y., Amano, T., Noda, M., et al. (2003). Alterations of structure and hydrolase activity of parkinsonism-associated human ubiquitin carboxyl-terminal hydrolase L1 variants. Biochem Biophys Res Commun *304*, 176-183. 10.1016/s0006-291x(03)00555-2.
15. Falb, R.J., Muller, A.J., Klein, W., Grimmel, M., Grasshoff, U., Spranger, S., Stobe, P., Gauck, D., Kuechler, A., Dikow, N., et al. (2021). Bi-allelic loss-of-function variants in KIF21A cause severe fetal akinesia with arthrogryposis multiplex. J Med Genet. 10.1136/jmedgenet-2021-108064.
16. Weisschuh, N., Mazzola, P., Heinrich, T., Haack, T., Wissinger, B., Tonagel, F., and Kelbsch, C. (2020). First submicroscopic inversion of the OPA1 gene identified in dominant optic atrophy - a case report. BMC Med Genet *21*, 236. 10.1186/s12881-020-01166-z.
17. Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W., Hegde, M., Lyon, E., Spector, E., 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. Genet Med *17*, 405-424. 10.1038/gim.2015.30.
18. Susak, H., Serra-Saurina, L., Demidov, G., Rabionet, R., Domenech, L., Bosio, M., Muyas, F., Estivill, X., Escaramis, G., and Ossowski, S. (2021). Efficient and flexible Integration of variant characteristics in rare variant association studies using integrated nested Laplace approximation. PLoS Comput Biol *17*, e1007784. 10.1371/journal.pcbi.1007784.
19. Investigators, G.P.P., Smedley, D., Smith, K.R., Martin, A., Thomas, E.A., McDonagh, E.M., Cipriani, V., Ellingford, J.M., Arno, G., Tucci, A., et al. (2021). 100,000 Genomes Pilot on Rare-Disease Diagnosis in Health Care - Preliminary Report. N Engl J Med *385*, 1868-1880. 10.1056/NEJMoa2035790.
20. Smedley, D., Jacobsen, J.O., Jager, M., Kohler, S., Holtgrewe, M., Schubach, M., Siragusa, E., Zemojtel, T., Buske, O.J., Washington, N.L., et al. (2015). Next-generation diagnostics and disease-gene discovery with the Exomiser. Nat Protoc *10*, 2004-2015. 10.1038/nprot.2015.124.
21. Havrilla, J.M., Pedersen, B.S., Layer, R.M., and Quinlan, A.R. (2019). A map of constrained coding regions in the human genome. Nat Genet *51*, 88-95. 10.1038/s41588-018-0294-6.
22. Martin, A.R., Williams, E., Foulger, R.E., Leigh, S., Daugherty, L.C., Niblock, O., Leong, I.U.S., Smith, K.R., Gerasimenko, O., Haraldsdottir, E., et al. (2019). PanelApp crowdsources expert knowledge to establish consensus diagnostic gene panels. Nat Genet *51*, 1560-1565. 10.1038/s41588-019-0528-2.
23. Dolzhenko, E., Deshpande, V., Schlesinger, F., Krusche, P., Petrovski, R., Chen, S., Emig-Agius, D., Gross, A., Narzisi, G., Bowman, B., et al. (2019). ExpansionHunter: a sequence-graph-based tool to analyze variation in short tandem repeat regions. Bioinformatics *35*, 4754-4756. 10.1093/bioinformatics/btz431.
24. Kohler, S., Gargano, M., Matentzoglu, N., Carmody, L.C., Lewis-Smith, D., Vasilevsky, N.A., Danis, D., Balagura, G., Baynam, G., Brower, A.M., et al. (2021). The Human Phenotype Ontology in 2021. Nucleic Acids Res *49*, D1207-D1217. 10.1093/nar/gkaa1043.
25. Hengel, H., Hannan, S.B., Dyack, S., MacKay, S.B., Schatz, U., Fleger, M., Kurringer, A., Balousha, G., Ghanim, Z., Alkuraya, F.S., et al. (2021). Bi-allelic loss-of-function variants in BCAS3 cause a syndromic neurodevelopmental disorder. Am J Hum Genet *108*, 1069-1082. 10.1016/j.ajhg.2021.04.024.

26. Parodi, L., Fenu, S., Barbier, M., Banneau, G., Duyckaerts, C., Tezenas du Montcel, S., Monin, M.L., Ait Said, S., Guegan, J., Tallaksen, C.M.E., et al. (2018). Spastic paraplegia due to SPAST mutations is modified by the underlying mutation and sex. Brain *141*, 3331-3342. 10.1093/brain/awy285.
27. Goizet, C., Boukhris, A., Mundwiller, E., Tallaksen, C., Forlani, S., Toutain, A., Carriere, N., Paquis, V., Depienne, C., Durr, A., et al. (2009). Complicated forms of autosomal dominant hereditary spastic paraplegia are frequent in SPG10. Human mutation *30*, E376-385. 10.1002/humu.20920.
28. Turner, M.R., Rabiner, E.A., Al-Chalabi, A., Shaw, C.E., Brooks, D.J., Leigh, P.N., and Andersen, P.M. (2007). Cortical 5-HT1A receptor binding in patients with homozygous D90A SOD1 vs sporadic ALS. Neurology *68*, 1233-1235. 10.1212/01.wnl.0000259083.31837.64.
29. McMacken, G., Lochmuller, H., Bansagi, B., Pyle, A., Lochmuller, A., Chinnery, P.F., Laurie, S., Beltran, S., Matalonga, L., and Horvath, R. (2020). Behr syndrome and hypertrophic cardiomyopathy in a family with a novel UCHL1 deletion. J Neurol *267*, 3643-3649. 10.1007/s00415-020-10059-3.
30. Chen, F., Sugiura, Y., Myers, K.G., Liu, Y., and Lin, W. (2010). Ubiquitin carboxyl-terminal hydrolase L1 is required for maintaining the structure and function of the neuromuscular junction. Proc Natl Acad Sci U S A *107*, 1636-1641. 10.1073/pnas.0911516107.
31. Auer-Grumbach, M., Toegel, S., Schabhüttl, M., Weinmann, D., Chiari, C., Bennett, D.L.H., Beetz, C., Klein, D., Andersen, P.M., Böhme, I., et al. (2016). Rare Variants in MME, Encoding Metalloprotease Neprilysin, Are Linked to Late-Onset Autosomal-Dominant Axonal Polyneuropathies. Am J Hum Genet *99*, 607-623. 10.1016/j.ajhg.2016.07.008.
32. Depondt, C., Donatello, S., Rai, M., Wang, F.C., Manto, M., Simonis, N., and Pandolfo, M. (2016). MME mutation in dominant spinocerebellar ataxia with neuropathy (SCA43). Neurol Genet *2*, e94. 10.1212/NXG.0000000000000094.
33. Guglielmotto, M., Monteleone, D., Vasciaveo, V., Repetto, I.E., Manassero, G., Tabaton, M., and Tamagno, E. (2017). The Decrease of Uch-L1 Activity Is a Common Mechanism Responsible for Aβ 42 Accumulation in Alzheimer's and Vascular Disease. Front Aging Neurosci *9*, 320. 10.3389/fnagi.2017.00320.
34. Maraganore, D.M., Lesnick, T.G., Elbaz, A., Chartier-Harlin, M.C., Gasser, T., Kruger, R., Hattori, N., Mellick, G.D., Quattrone, A., Satoh, J., et al. (2004). UCHL1 is a Parkinson's disease susceptibility gene. Ann Neurol *55*, 512-521. 10.1002/ana.20017.
35. Healy, D.G., Abou-Sleiman, P.M., Casas, J.P., Ahmadi, K.R., Lynch, T., Gandhi, S., Muqit, M.M., Foltynie, T., Barker, R., Bhatia, K.P., et al. (2006). UCHL-1 is not a Parkinson's disease susceptibility gene. Ann Neurol *59*, 627-633. 10.1002/ana.20757.
36. Zhang, M., Cai, F., Zhang, S., Zhang, S., and Song, W. (2014). Overexpression of ubiquitin carboxyl-terminal hydrolase L1 (UCHL1) delays Alzheimer's progression in vivo. Sci Rep *4*, 7298. 10.1038/srep07298.
37. Gong, B., Cao, Z., Zheng, P., Vitolo, O.V., Liu, S., Staniszewski, A., Moolman, D., Zhang, H., Shelanski, M., and Arancio, O. (2006). Ubiquitin hydrolase Uch-L1 rescues beta-amyloid-induced decreases in synaptic function and contextual memory. Cell *126*, 775-788. 10.1016/j.cell.2006.06.046.
38. Oeckl, P., Weydt, P., Thal, D.R., Weishaupt, J.H., Ludolph, A.C., and Otto, M. (2020). Proteomics in cerebrospinal fluid and spinal cord suggests UCHL1, MAP2 and GPNMB as biomarkers and underpins importance of transcriptional pathways in amyotrophic lateral sclerosis. Acta Neuropathol *139*, 119-134. 10.1007/s00401-019-02093-x.
39. Carrieri, C., Cimatti, L., Biagioli, M., Beugnet, A., Zucchelli, S., Fedele, S., Pesce, E., Ferrer, I., Collavin, L., Santoro, C., et al. (2012). Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. Nature *491*, 454-457. 10.1038/nature11508.
40. Zucchelli, S., Fasolo, F., Russo, R., Cimatti, L., Patrucco, L., Takahashi, H., Jones, M.H., Santoro, C., Sblattero, D., Cotella, D., et al. (2015). SINEUPs are modular antisense long non-coding

RNAs that increase synthesis of target proteins in cells. Front Cell Neurosci *9*, 174. 10.3389/fncel.2015.00174.

FIGURE TITLES AND LEGENDS

Figure 1: Pedigrees of 17 *UCHL1* Families, and Structure of *UCHL1*

(a) Pedigrees of families segregating heterozygous LoFs in *UCHL1* (F1-F14), and one *UCHL1* inframe duplication (F15-F17). Affected family members are shown in closed symbols and healthy family members with open symbols. wt stands for wildtype and p for pathogenic variant. Other family members were not available for genetic testing.

(b) Graphical illustration of *UCHL1* gene and protein domain structures, and position of the identified variants. Newly reported LoFs are written in red and previously published variants in black.

Figure 2: Distribution of symptoms in *UCHL1* patients

Orange bars represent the proportion of patients showing given clinical features and carrying dominant-acting heterozygous *UCHL1* variants. Grey bars illustrate the percentage of SPG79 patients carrying bi-allelic *UCHL1* variants.[7-9,29] The symptoms are defined using the Human Phenotype Ontology (HPO) Terms.

Figure 3: Mass spectrometry based proteomics

(a) Volcano plot resulting from nanoLC-MS/MS comparing five patient fibroblast cell lines to 5 control fibroblast cells lines, both in duplicates. Statistically significant overexpressed proteins are depicted in red, significant underexpressed proteins are depicted in blue.

(b) Interactome based on most over-or underexpressed proteins. Interaction network computed using string-db, showing the 30 most over-or underexpressed proteins from nanoLC-MS/MS data. Overexpressed proteins are shown in red, underexpressed proteins in blue. Significant enrichment of interactions among these proteins is present (PPI 1.17e-7).