BRAIN
ORIGINAL ARTICLE

Combined genomics and proteomics unveils elusive variants and vast aetiologic heterogeneity in dystonia

Michael Zech,[1,2,3,4,†] Ivana Dzinovic,[1,2,†] Matej Skorvanek,[5,6,†] Philip Harrer,[1,2,†] Jan Necpal,[7,8] Robert Kopajtich,[1,2] Volker Kittke,[1,2] Erik Tilch,[1,2] Chen Zhao,[1,2] Eugenia Tsoma,[9] Ugo Sorrentino,[1,2,10] Elisabetta Indelicato,[1,2,11] Antonia Stehr,[1] Alice Saparov,[1,2] Lucia Abela,[12] Miriam Adamovicova,[13] Alexandra Afenjar,[14] Birgit Assmann,[15] Janette Baloghova,[16,17] Matthias Baumann,[18] Riccardo Berutti,[1,2,19] Zuzana Brezna,[8] Melanie Brugger,[1,4,20] Theresa Brunet,[1,4,21] Benjamin Cogne,[22,23] Isabel Colangelo,[24] Erin Conboy,[25] Felix Distelmaier,[26,27] Matthias Eckenweiler,[28] Barbara Garavaglia,[24] Arie Geerlof,[29] Elisabeth Graf,[1,4] Annette Hackenberg,[12] Denisa Harvanova,[30] Bernhard Haslinger,[31] Petra Havrankova,[32] Georg F. Hoffmann,[15] Wibke G. Janzarik,[28] Boris Keren,[33] Miriam Kolnikova,[34] Konstantinos Kolokotronis,[35] Zuzana Kosutzka,[36] Anne Koy,[37,38] Martin Krenn,[39,40] Magdalena Krygier,[41] Katarina Kusikova,[34] Oliver Maier,[42] Thomas Meitinger,[1,4] Christian Mertes,[1,4,43,44] Ivan Milenkovic,[39] Edoardo Monfrini,[45,46] Andre Santos Dias Mourao,[29] Thomas Musacchio,[47,48] Mathilde Nizon,[23] Miriam Ostrozovicova,[5,6] Martin Pavlov,[1,2] Iva Prihodova,[32] Irena Rektorova,[49,50] Luigi M. Romito,[51] Barbora Rybanska,[34] Ariane Sadr-Nabavi,[1,2,52,53,54,55] Susanne Schwenger,[56] Ali Shoeibi,[54] Alexandra Sitzberger,[21] Dmitrii Smirnov,[1,2] Jana Svantnerova,[36] Raushana Tautanova,[57] Sandra P. Toelle,[12] Olga Ulmanova,[32] Francesco Vetrini,[25] Katharina Vill,[21] Matias Wagner,[1,2,4,21] David Weise,[58,59] Giovanna Zorzi,[60] Alessio Di Fonzo,[45,46] Konrad Oexle,[1,2] Steffen Berweck,[21,61] Volker Mall,[56,62] Sylvia Boesch,[11] Barbara Schormair,[1,2,‡] Holger Prokisch,[1,2,63,‡] Robert Jech[32,‡] and Juliane Winkelmann[1,2,4,64,65,‡]

[†,‡]These authors contributed equally to this work.

Dystonia is a rare disease trait for which large-scale genomic investigations are still underrepresented. Genetic heterogeneity among patients with unexplained dystonia warrants interrogation of entire genome sequences, but this has not yet been systematically evaluated.

To significantly enhance our understanding of the genetic contribution to dystonia, we (re)analysed 2874 whole-exome sequencing (WES), 564 whole-genome sequencing (WGS), as well as 80 fibroblast-derived proteomics

Received October 04, 2024. Revised January 08, 2025. Accepted January 26, 2025. Advance access publication February 12, 2025

datasets, representing the output of high-throughput analyses in 1990 patients and 973 unaffected relatives from 1877 families. Recruitment and precision-phenotyping procedures were driven by long-term collaborations of international experts with access to overlooked populations.

By exploring WES data, we found that continuous scaling of sample sizes resulted in steady gains in the number of associated disease genes without plateauing. On average, every second diagnosis involved a gene not previously implicated in our cohort. Second-line WGS focused on a subcohort of undiagnosed individuals with high likelihood of having monogenic forms of dystonia, comprising large proportions of patients with early onset (81.3%), generalized symptom distribution (50.8%) and/or coexisting features (68.9%). We undertook extensive searches for variants in nuclear and mitochondrial genomes to uncover 38 (ultra)rare diagnostic-grade findings in 37 of 305 index patients (12.1%), many of which had remained undetected due to methodological inferiority of WES or pipeline limitations. WGS-identified elusive variations included alterations in exons poorly covered by WES, RNA-gene variants, mitochondrial-DNA mutations, small copy-number variants, complex rearranged genome structure and short tandem repeats. For improved variant interpretation in WGS-inconclusive cases, we employed systematic integration of quantitative proteomics. This aided in verifying diagnoses related to technically challenging variants and in upgrading a variant of uncertain significance (3 of 70 WGS-inconclusive index patients, 4.3%). Further, unsupervised proteomic outlier analysis supplemented with transcriptome sequencing revealed pathological gene underexpression induced by transcript disruptions in three more index patients with underlying (deep) intronic variants (3/70, 4.3%), highlighting the potential for targeted antisense-oligonucleotide therapy development. Finally, trio-WGS prioritized a *de novo* missense change in the candidate *PRMT1*, encoding a histone methyltransferase. Data-sharing strategies supported the discovery of three distinct *PRMT1 de novo* variants in four phenotypically similar patients, associated with loss-of-function effects in *in vitro* assays.

This work underscores the importance of continually expanding sequencing cohorts to characterize the extensive spectrum of gene aberrations in dystonia. We show that a pool of unresolved cases is amenable to WGS and complementary multi-omic studies, directing advanced aetiopathological concepts and future diagnostic-practice workflows for dystonia.

1 Institute of Human Genetics, School of Medicine and Health, Technical University of Munich, Munich 81675, Germany
2 Institute of Neurogenomics, Helmholtz Zentrum München, Munich 85764, Germany
3 Institute for Advanced Study, Technical University of Munich, Garching 85748, Germany
4 Bavarian Genomes Network for Rare Disorders, Munich 81675, Germany
5 Department of Neurology, P.J. Safarik University, Kosice 4001, Slovakia
6 Department of Neurology, University Hospital of L. Pasteur, Kosice 4001, Slovakia
7 Department of Neurology, Zvolen Hospital, Zvolen 96001, Slovakia
8 Parkinsonism and Movement Disorders Treatment Center, Zvolen Hospital, Zvolen 96001, Slovakia
9 Department of Family Medicine and Outpatient Care, Regional Clinical Center of Neurosurgery and Neurology, Uzhhorod National University, Uzhhorod 88000, Ukraine
10 Department of Women's and Children's Health, Clinical Genetics Unit, University of Padova, Padua 35128, Italy
11 Department of Neurology, Center for Rare Movement Disorders Innsbruck, Medical University of Innsbruck, Innsbruck 6020, Austria
12 Department of Pediatric Neurology, University Children's Hospital Zürich, University of Zürich, Zurich 8008, Switzerland
13 Department of Paediatric Neurology, Thomayer University Hospital, Prague 12108, Czech Republic
14 Clinical Genetics Unit, Reference Center for Cerebellar Malformations and Congenital Diseases, Armand-Trousseau Hospital, APHP, Sorbonne University, Paris 75013, France
15 Center for Pediatric and Adolescent Medicine, Clinic I, University Hospital Heidelberg, Heidelberg University, Heidelberg 69120, Germany
16 Faculty of Medicine, Department of Dermatovenerology, P. J. Safarik University, Kosice 4001, Slovakia
17 Department of Dermatovenerology, University Hospital of L. Pasteur, Kosice 4001, Slovakia
18 Division of Pediatric Neurology, Department of Pediatrics I, Medical University of Innsbruck, Innsbruck 6020, Austria
19 Center for Advanced Studies, Research and Development in Sardinia, NGS Bioinformatics, Pula 09010, Italy
20 Department of Obstetrics and Gynecology, Klinikum Rechts der Isar, Technical University of Munich, Munich 81675, Germany
21 Division of Pediatric Neurology and Developmental Medicine and LMU Center for Children with Medical Complexity, Dr. von Hauner Children's Hospital, LMU Hospital, Ludwig-Maximilians-Universität, Munich 80337, Germany

22 Nantes Université, CHU de Nantes, CNRS, INSERM, l'institut du thorax, Nantes 44000, France
23 Nantes Université, CHU de Nantes, Service de Génétique médicale, Nantes 44000, France
24 Medical Genetics and Neurogenetics Unit, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan 20126, Italy
25 Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN 46202, USA
26 Department of General Pediatrics, Neonatology, and Pediatric Cardiology, Medical Faculty and University Hospital Düsseldorf, Heinrich-Heine-University, Düsseldorf 40225, Germany
27 Department of General Pediatrics, West German Center for Child and Adolescent Health, Düsseldorf 40225, Germany
28 Faculty of Medicine, Department of Neuropediatrics and Muscle Disorders, University Medical Center, University of Freiburg, Freiburg 79110, Germany
29 Institute of Structural Biology, Helmholtz Center Munich, Munich 85764, Germany
30 Faculty of Medicine, Associated Tissue Bank, P. J. Safarik University and L. Pasteur University Hospital in Kosice, Kosice 4001, Slovakia
31 Department of Neurology, Klinikum rechts der Isar, School of Medicine and Health, Technical University of Munich, Munich 81675, Germany
32 Department of Neurology and Center of Clinical Neuroscience, First Faculty of Medicine Charles University and General University Hospital in Prague, Prague 12108, Czech Republic
33 Department of Genetics, Pitié-Salpêtrière Hospital, APHP, Sorbonne University, Paris 75013, France
34 Faculty of Medicine, Department of Pediatric Neurology, Comenius University, University Hospital Bratislava and National Institute of Children's Diseases, Bratislava 83340, Slovakia
35 Institute of Medical Genetics, University of Zurich, Zurich 8952, Switzerland
36 Faculty of Medicine, Second Department of Neurology, Comenius University, University Hospital Bratislava, Bratislava 83340, Slovakia
37 Faculty of Medicine and University Hospital Cologne, Department of Pediatrics, University of Cologne, Cologne 50937, Germany
38 Faculty of Medicine and University Hospital Cologne, Center for Rare Diseases, University of Cologne, Cologne 50937, Germany
39 Department of Neurology, Medical University of Vienna, Vienna 1090, Austria
40 Comprehensive Center for Clinical Neurosciences & Mental Health, Medical University of Vienna, Vienna 1090, Austria
41 Department of Developmental Neurology, Medical University of Gdansk, Gdansk 80210, Poland
42 Division of Child Neurology, Department of Pediatrics, Children's Hospital, St. Gallen 9000, Switzerland
43 School of Computation, Information and Technology, Technical University of Munich, Garching 85748, Germany
44 Munich Data Science Institute, Technical University of Munich, Garching 85748, Germany
45 Department of Pathophysiology and Transplantation, Dino Ferrari Center, Neuroscience Section, University of Milan, Milan 20122, Italy
46 Foundation IRCCS Ca' Granda Ospedale Maggiore Policlinico, Neurology Unit, Milan 20122, Italy
47 Department of Neurology, University Hospital of Würzburg, Wurzburg 97080, Germany
48 Center for Rare Diseases, University Hospital of Würzburg, Wurzburg 97080, Germany
49 Brain and Mind Research Program, CEITEC, Masaryk University, Brno 62500, Czech Republic
50 First Department of Neurology, St. Anne's University Hospital and Faculty of Medicine, Masaryk University, Brno 62500, Czech Republic
51 Parkinson and Movement Disorders Unit, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan 20126, Italy
52 Faculty of Medicine, Department of Medical Genetics, Mashhad University of Medical Sciences, Mashhad 9177948564, Iran
53 Medical Genetics Research Center, Mashhad University of Medical Sciences, Mashhad 9177948564, Iran
54 Faculty of Medicine, Department of Neurology, Mashhad University of Medical Sciences, Qaem Medical Center, Mashhad 9177948564, Iran
55 Academic Center for Education, Culture and Research (ACECR)-Khorasan Razavi, Mashhad 9177948564, Iran
56 kbo-Kinderzentrum München, Munich 81377, Germany
57 Department of Neurosurgery, Medical Centre Hospital of the President's Affairs Administration of the Republic of Kazakhstan, Astana E495, Kazakhstan
58 Department of Neurology, Asklepios Fachklinikum Stadtroda, Stadtroda 07646, Germany
59 Department of Neurology, University of Leipzig, Leipzig 04103, Germany
60 Department of Pediatric Neuroscience, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan 20126, Italy
61 Specialist Center for Pediatric Neurology, Neurorehabilitation and Epileptology, Schön Clinic, Vogtareuth 83569, Germany
62 Chair of Social Pediatrics, Technical University of Munich, Munich 80333, Germany
63 DZKJ, Deutsches Zentrum Für Kinder- und Jugendgesundheit, Munich 80333, Germany
64 DZPG, Deutsches Zentrum Für Psychische Gesundheit, Munich 80333, Germany
65 Munich Cluster for Systems Neurology, SyNergy, Munich 81377, Germany

Correspondence to: PD Dr Michael Zech
Institute of Neurogenomics, Helmholtz Zentrum München, Deutsches Forschungszentrum für Gesundheit und Umwelt (GmbH), Ingolstädter Landstraße 1, Neuherberg 85764, Germany
E-mail: michael.zech@mri.tum.de

Keywords: dystonia; genomics; whole-genome sequencing; multi-omics; proteomics; transcriptomics

Introduction

The presentation of dystonia includes a broad diversity of clinical characteristics, ranging from isolated expressions affecting only part of the body to generalized manifestations with or without alternative clinical features.[1,2] Dystonia can be diagnosed from the neonatal period to late adulthood, requiring involvement of multiple specialists.[3,4] Our incomplete knowledge about the aetiologic subtypes of dystonia limits the potential to stratify patients for clinical trials and therapy development.[5,6]

Panel and whole-exome sequencing (WES) analyses in single patients, small case series and a few larger populations have shown that around 500 genes can be associated with monogenic dystonia.[7-14] Notwithstanding this, the steady gains in disease-gene discovery by systematic stepwise addition of newly sequenced individuals within a defined cohort have not been well documented.[15,16] Even among the most severe manifestations (e.g. early-onset generalized dystonia with additional congenital anomalies), the maximum molecular diagnosis rate of WES is still capped at around 50%.[17] Whole-genome sequencing (WGS) has the added benefit of detecting all possibly disease-relevant variations, but widespread application requires demonstration of its diagnostic efficacy in individual indications.[18-20] The role of WGS as a second-tier test in dystonia-affected patients with uninformative WES is unknown.[21] Furthermore, many available WGS studies do not provide the capacity for high-throughput functional annotation of variants and downstream mutational effect-guided gene prioritization.[18,19,22] Quantitative proteomics paired with other multi-omic profiling methods has been associated with improved diagnostic performance in selected genome-sequenced individuals with rare metabolic disease phenotypes.[23-25] By contrast, multilayered molecular testing strategies are understudied in dystonia.[15,26] Compared with some multi-omic modalities such as transcriptomics, the integration of proteomics in variant-analysis workflows has lagged behind.[20,27-29]

Over a decade, we have established a network involving multi-skilled experts, patient organizations and specialist referral sites to provide a systematic approach to the diagnosis of genetic dystonia.[30] In this work, we demonstrate scalability of WES analyses for continuously recruited, geographically diverse groups of dystonia-affected individuals with linear increase in gene identification. We report transition from WES to WGS in a population of long-term undiagnosed cases with more severe forms of the disease to expand diagnostic capabilities and assess disregarded variants including mutations cryptic to prior testing. We genotyped short tandem repeats (STRs),[31] noting an unexpected role of repeat-expansion disorders in dystonia. Finally, we increased the aetiologic yield of WGS by cohortwide proteomics and an array of case-specific analysis strategies including RNA sequencing (RNA-seq), functional assays and data-sharing, the latter two of which contributed to delineation of a previously undescribed monogenic disease with dystonia due to *PRMT1* variants.

Our study underscores the imperative for genome-wide analysis of all variant types and implementation of multifaceted adjunct functional tests to augment diagnostic rates in rare dystonias.

Materials and methods

Patients and study design

The index patients and additional family members were recruited through multi-site research collaboration from 38 institutions in 11 countries between 2015 and 2024[7,8,30] (Fig. 1 and Supplementary Table 1). Informed consent for research studies was obtained for all participants in accordance with the ethical guidelines of our institutional review boards. Patients with secondary or known monogenic dystonia were ineligible. Standardized data on clinical manifestations and long-term outcomes including videoed examinations were individually curated for use in variant prioritization and interpretation of genomic findings. The phenotyping entailed a comprehensive documentation of signs and symptoms, examined by established neurologic assessment schemes and rating recommendations,[1-4] as well as a thorough review of diagnosed comorbidities and available routine diagnostic test results. To enlarge the referral pool of patients from usually underrepresented geographical regions,[32] a network of investigators in Ukraine and Slovakia (the latter focusing on underserved minority groups) has been involved since 2022. A subset of the present cohort has been described in previous publications.[7,8,17] WES was the primary genetic investigation of all patients in this study, followed by WGS in a subset of individuals whose conditions remained unexplained after completion of exome-wide analysis (Fig. 1A). We prioritized cases for WGS based on the following criteria: (i) ongoing suspicion of an underlying monogenic disorder (necessary inclusion criterion), as defined by one or more of the following points:[17] dystonia onset <21 years, non-focal dystonia, coexisting features, ≥3 affected family members, referring clinician formulated a specific genetic differential diagnosis (e.g. on the basis of MRI abnormalities); (ii) relatives available for family-based WGS; and/or (iii) skin-biopsy sample obtainable for complementary multi-omic testing. There were two subgroups entering the WGS pathway: (i) patients who had received unrevealing WES by us under research protocols [253/305 WGS-cohort index patients (83.0%), individuals from the here-described WES studies]; and (ii) patients with 'negative' WES testing from external fee-for-service projects [52/305 WGS-cohort index patients (17.0%)], making our population representative of a wider spectrum of challenging cases seen in different laboratories (Fig. 1A). According to the study workflow, we did not re-analyse existing WES data of patients following their ascertainment for WGS.[25,33]

Whole-exome sequencing

Study participants underwent WES on Illumina systems over a span of time. Research paired-end sequencing and data analyses were performed according to standards of our accredited laboratory, as described previously.[7,8] There were 1164 newly recruited individuals [725 index patients, 2021–2024, 40.5% (1164/2874) of the total cohort] who had WES as part of the present study. Only likely pathogenic and pathogenic variants[34,35] in line with the expected mode of inheritance and related to the phenotype were retained for downstream analysis (Supplementary Table 2). Additionally,

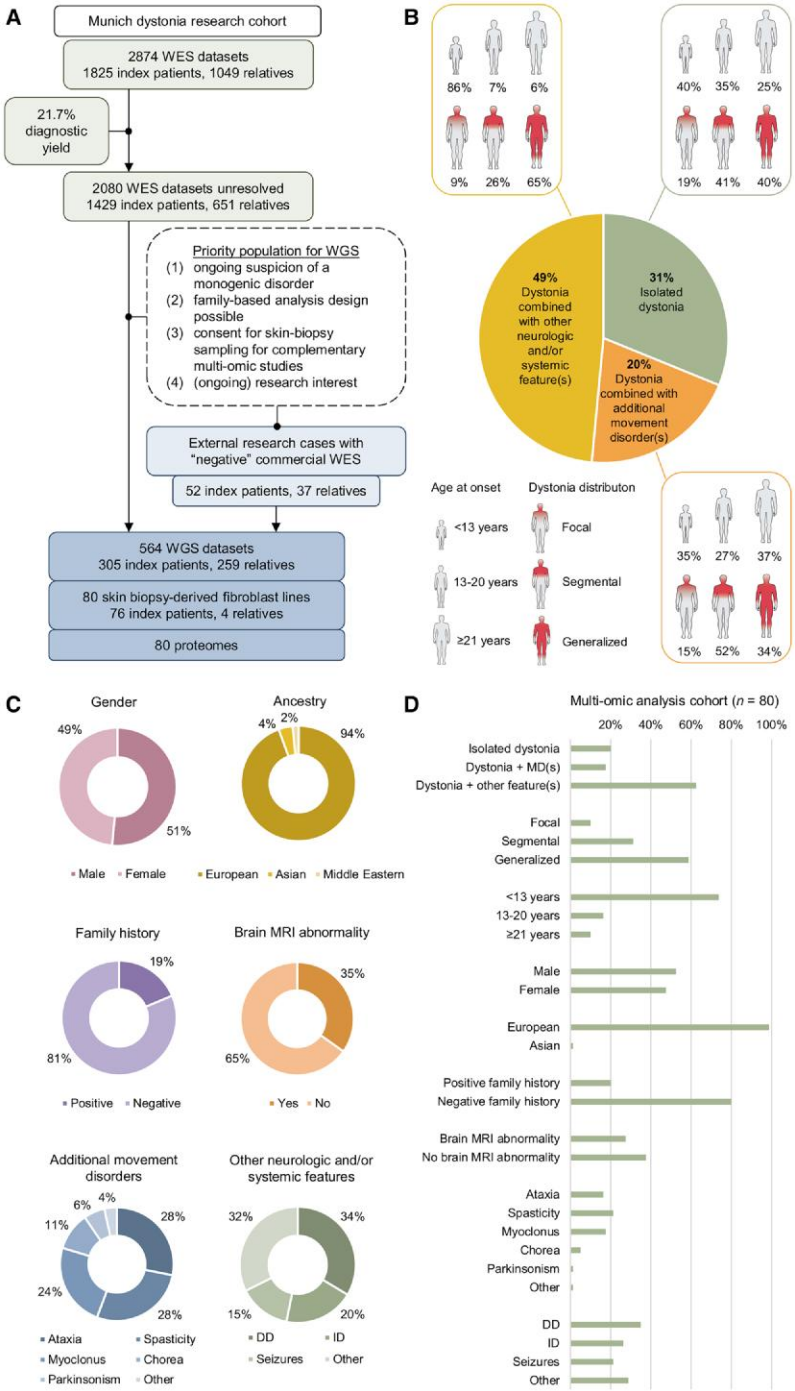

Figure 1 Study overview and overall cohort structures. (**A**) The WES cohort comprised 2874 individuals from 1825 families, analysed from 2015 to 2024. A total of 21.7% (396/1825) of the index patients received a molecular diagnosis, leaving 1429 unresolved families after exome-wide testing. Second-tier WGS was offered to a subset of undiagnosed cases and relatives, by applying inclusion criteria as follows: (1) ongoing suspicion of a monogenic condition, as defined by one or more of the following points[17]: dystonia onset <21 years, non-focal dystonia, coexisting features, ≥3 affected family members, referring clinician formulated a specific genetic differential diagnosis (e.g. on the basis of MRI abnormalities); (2) family-based sequencing design possible; and (3) available consent for skin-biopsy sampling to complete multi-omic studies. Criteria (2) and (3) were not fulfilled by all included patients. An additional 52 index patients with uninformative WES were ascertained for the WGS experiments from external laboratories following research consent. The overall WGS sample size was determined by funding capacity, totalling 564 individuals from 305 families. We performed quantitative proteomic analyses in fibroblasts from 80 patients who underwent WGS. (**B**) Presence or absence of coexisting features, age at dystonia onset and dystonia distribution for 305 index patients in the WGS cohort. (**C**) Summary of WGS-cohort index patients with regard to gender, ancestry, family history information, neuroimaging and comorbidities. Brain MRI data were available for 204 index patients (66.9%). One or more additional movement disorders were present in 146 index patients (47.9%), and other neurologic and/or systematic features in 148 index patients (48.5%); the distributions of non-dystonic movement symptoms and other features are shown in the *bottom* pie charts. (**D**) Summary of clinical characteristics for 80 patients (76 index patients and 4 affected relatives) with available fibroblast-derived proteomic datasets. Brain MRI data were available for 52 patients (65.0%). DD = developmental delay; ID = intellectual disability; MD = movement disorder; WES = whole-exome sequencing; WGS = whole-genome sequencing.

we applied an established burden testing scheme to the WES data of our cohort in order to discover genes with significant excess of rare predicted loss-of-function (pLoF) variants in patients compared with controls.[36,37]

Whole-genome sequencing

We initiated WGS to clinical standards for a prospective tightly ascertained group of WES-unresolved patients with dystonia in 2020. Blood-derived DNA of affected individuals and their healthy relatives was subjected to library preparation using TruSeq PCR-free kits (Illumina),[18,33] followed by paired-end 150-bp sequencing on NovaSeq6000 instruments at the next-generation sequencing Core Facility of University Hospital Bonn (Bonn, Germany), Helmholtz, Munich, or the Institute of Human Genetics of Technical University of Munich (Munich, Germany). The data generated per sample were >100 Gb, achieving an average depth of coverage of >40x with >20x coverage for at least 95% of the nuclear genome. Mitochondrial DNA was entirely covered with >1000x (Supplementary Table 3). The raw-data output was bioinformatically processed with an in-house-generated expert system that allowed for comprehensive variant analysis, EVAdb: https://github.com/mri-ihg/EVAdb. This platform integrated Burrows–Wheeler Aligner for genomic mapping (hg19) and the Genome Analysis Toolkit (GATK) for detection of single-nucleotide variants (SNVs), short insertions/deletions (indels) and mitochondrial (MT) variants. Discovery of structural variants (SVs) including copy-number variants (CNVs) was done using a combination of six algorithms: BreakDancer, CNVnator, LUMPY, Manta, Pindel and Whamg[38]; CNVs/SVs supported by ≥2 callers had higher analytic validity. Moreover, ExpansionHunter (EH) was deployed with default parameters in order to genotype STRs at 29 targeted loci of known relevance for neurologic disorders.[31,39] All datasets underwent careful filtration and phenotype-driven variant prioritization processes. First, we focused on known disease genes and looked for variants in genes that had previously been associated with features observed in our patients.[13] To that end, we obtained gene lists from the Online Mendelian Inheritance in Man (OMIM) database and PubMed searches, with internal curation as previously described.[7,8] Priority was given to genes most strongly linked to the phenotypes, but expanded gene–disease relationships were also considered. Commonly used strategies were applied to narrow down variants according to rarity, mode of segregation, consequence, prediction of deleteriousness and documented clinical significance.[20] Data from in-house-sequenced individuals, the Genome Aggregation Database (gnomAD) v.4.1.0[40,41] and other online repositories[13,42-45] were used to determine variant allele frequencies for patients of European and non-European ancestries. At this stage, non-coding variants were only taken into consideration if they had been reported before in ClinVar[45] or the Human Gene Mutation Database (HGMD).[46] MT variants were extracted based on overlap with confirmed pathogenic mutations in MITOMAP.[44] Prioritized CNVs/SVs had to satisfy the following requirements: (i) rare in 2000 in-house control genomes and reference databases[40,42,43]; (ii) interrupting a gene present on our disease–gene lists; and (iii) zygosity consistent with the known transmission of the disorder linked to the interrupted gene.[13] We additionally searched for STRs[31] that were expanded by comparing both allele sizes to locus-specific thresholds[47] and evaluated results in the context of the patients' clinical presentations. All datasets of patients with available skin biopsy-derived fibroblasts were further analysed in light of identified proteomic alterations (see below), facilitating the assessment of variants with difficult interpretability including non-coding variants.[48] Pathogenicity was determined according to appropriate American College of Medical Genetics and Genomics (ACMG) interpretation standards.[34,35] Second, for all undiagnosed cases, we filtered for candidate variants in genes without established disease association; this strategy focused on damaging *de novo* heterozygous variants in mutation-intolerant genes and expected high-impact recessive alterations (i.e. bi-allelic gene-disrupting variants).[20,49] Thereafter, a review of each case was performed to reach multidisciplinary consensus that proposed variants were deemed causative for the phenotype or plausible candidate aetiologies requiring further investigation. We visually inspected the read evidence for all SNVs, indels, MT variants and CNVs/SVs selected for reporting as diagnostic or candidate findings by using the Integrative Genomics Viewer (IGV).[50]

Orthogonal methods

WGS-identified STR expansions in *FXN*, *HTT* and *PABPN1* were validated by standard PCR-based techniques in accredited test laboratories. Additional evidence supportive of a causative association with identified variants was generated by the best applicable validated methods such as analysis of alpha-fetoprotein (AFP) concentration[51] (*ATM*-mutated patients), plasma amino-acid profiling[52] (*GLS*-mutated patient) and thyroid function testing[53] (*SLC16A2*-mutated female patient).

Immunoblot analysis in patient fibroblasts

Fibroblast homogenates were produced according to described methodologies.[30] Total protein lysates processed by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis were probed with the following primary antibodies: ANK2 (Santa Cruz, sc-12718, 1:1000), GLS (GeneTex, GTX131263, 1:5000), GAPDH (Sigma, G8795-25UL, 1:20 000) and β-tubulin (abeomics, 11-13002, 1:5000). Blots were incubated with species-appropriate secondary antibodies (anti-rabbit, Biolegend, 406401, 1:1000; anti-mouse, Jackson, 115-036-062, 1:10 000).

Glutaminase activity assay

We assayed the activity of the *GLS*-encoded enzyme glutaminase in patient and control fibroblasts by using the Glutaminase (GLS) Activity Assay Kit (Elabscience, E-BC-K660-M) per the manufacturer's specifications.

RNA sequencing

Transcriptome libraries were created from fibroblast-derived mRNA using the same strand-specific, polyA-tailed protocol (Illumina TruSeq) as in our previous RNA-seq studies.[54,55] All the mRNA samples used in the experiments had a high quality score (RNA integrity number of 10 for each sample). Each RNA-seq assay achieved a median sequencing depth of at least 50 million reads,[55] with a high percentage of accurately aligned reads (>80% for all samples).[55] RNA-seq data were assessed on a per patient basis, allowing us to examine unique transcriptional perturbations related to genomic variants.[54] WGS-prioritized mutational events with potential impact on splicing were manually evaluated using IGV.[54] For investigation of outliers of gene expression, we ran the Outlier in RNA-Seq Finder (OUTRIDER) tool using the recommended default settings.[55,56] Fibroblast RNA-seq data of 269 individuals with non-dystonic Mendelian conditions were used as controls.[55,57]

Quantitative proteomic analysis

We augmented our WGS experiments with the study of overall protein signatures of patient fibroblasts to explore the added value of proteomics for diagnosis in dystonia. The proteomics-guided framework was based on the evaluation of protein expression changes that were unique to the specific patient, as outlined in our earlier work.[23] We used proteomic profiling as a complementary diagnostic tool to WGS in two distinct ways: (i) we initiated the analytic protocol from observed genomic variants of indeterminate diagnostic confidence with the goal of validating them as relevant contributors to the phenotype; and (ii) we utilized the datasets of cases without candidates to search globally for outliers of expression and nominate genes for closer analysis. Detailed assay protocols of liquid chromatography-mass spectrometry (LC-MS) on fibroblast extracts and the raw-file processing strategies have been reported elsewhere.[23,54,57,58] LC-MS data were acquired at the BayBioMS core facility of the Technical University of Munich (Freising, Germany) on a Fusion Lumos Tribrid mass spectrometer (Thermo Scientific). Outputs were analysed using the MaxQuant platform,[59] identifying peptides and reviewed canonical and isoform proteins.[23] Analyses were performed in separate batches over the course of the study. For quantitative ranking and outlier detection, we used in-house developed pipelines and software including PROTRIDER (https://github.com/gagneurlab/py_outrider),[23,57] which provided a list of expression outliers for each sample along with information on multiple testing-corrected *P*-values, z-scores and fold-change (FC) for deviations compared with the proteomics test cohort. PROTRIDER, a denoising autoencoder-based algorithm,[56] was designed to automatically correct for known and unknown confounders including batch effects and other technical variables.[23] Our large total cohort ($n = 440$ proteomes; 80 from patients with dystonia, 360 from independent in-house research participants with rare disorders)[57] served as an effective control dataset for investigation of underexpression and outlier status in each individual case subject. Proteins that were not detected in ≥50% of samples were removed from the analysis. We specifically valued all underexpressed proteins with adjusted $P < 0.05$,[23] but also extended filtering parameters to scrutinize other lower-expressed products of dystonia-associated genes[7,8,13] in patients of interest.

Functional characterization of candidate *PRMT1* variants

To test the functional impact of *PRMT1* missense changes, we designed gBlock DNA sequences (Integrated DNA Technologies) of normal control and mutant protein arginine methyltransferase 1 (PRMT1). Four mutants were created to carry single-residue substitutions seen in patients from this study ($n = 4$) or the literature ($n = 1$),[60] and one to carry a gnomAD-listed 'control' variant.[40] DNA fragments were inserted into the pETM-11 bacterial expression vector[61] containing an N-terminal 6xHis-Tag by restriction enzyme cloning. Successful integration was confirmed by Sanger sequencing. Vectors were transformed into the *Escherichia coli* strain Rosetta2 (DE3) pLysS for overexpression. Expression levels of normal control and mutant versions of PRMT1 in bacterial lysates were analysed by immunoblotting using anti-Prmt1 (Abcam, ab73246); anti-RecA (Abcam, ab63797) served as the internal control. PRMT1 proteins were purified as previously described,[62] and methyltransferase activity was measured using the MTase-Glo™ Methyltransferase Assay (Promega).

Results

Multinational dystonia cohort for genomic analyses

A total of 2874 individuals (1933 patients and 941 healthy family members) were enrolled by us for WES (Fig. 1A and Supplementary Table 1). Among index patients, females were slightly overrepresented (955/1825, 52.3%). The participants were from 1825 families of diverse geographical origin. We actively included families from countries that appeared underrepresented in previous dystonia sequencing projects,[7,11,15] focusing on these populations especially during the new recruitment period from 2021 to 2024 (Fig. 2A). We performed WES in more than one affected member in 538 families (55 duos, 450 trios, 28 quartets, 5 multiplex pedigrees).

Review of our WES cohort identified 253 families who met eligibility criteria for WGS and were interested in this follow-up. Additionally, 52 families who did not have an identified molecular basis after external WES were enrolled for WGS (Fig. 1A). In total, 177 WGS-eligible patients (177/305, 58.0%) were recruited as singletons (index-only), while for 128 cases (128/305, 42.0%) at least one additional family member was ascertained for WGS (5 duos, 111 trios, 8 quartets, 4 multiplex pedigrees). Demographic and clinical features of the WGS-analysed index patients are summarized in Fig. 1B and C; the majority of cases had early disease onset (<21 years, 81.3%; 61.3% with onset in infancy or childhood) and displayed generalized or segmental dystonia (50.8% and 35.7%, respectively) (Fig. 1B). Coexisting symptoms were present in 68.9% of patients, comprising a wide spectrum of abnormalities including additional movement disorders, developmental delay, intellectual disability and other features (Fig. 1C).

Of the patients who proceeded to WGS, 80 (76 index patients; 76/305, 24.9%) consented for skin-biopsy sampling. Figure 1D shows an overview of the characteristics of patients with fibroblasts available for multi-omic analyses. Proteomics was completed for all biopsied patients (Fig. 1A), whereas RNA-seq was done in a subset ($n = 9$) for whom this method was expected to help further with evaluation of specific variants.

Longitudinal evaluation of reportable whole-exome sequencing findings

Of the 1825 index patients tested by WES, 396 (21.7%) were molecularly diagnosed over the course of the study (Fig. 2A and B and Supplementary Table 2). Most diagnoses came from likely pathogenic/pathogenic[34] coding and splice-site SNVs (298/404; 73.8%) or indels (103/404; 25.5%), whereas screening for likely pathogenic/pathogenic[35] exonic deletion CNVs and larger microdeletions yielded another 6.4% (26/404) of diagnoses. We were significantly more successful in establishing diagnoses in patients who had at least one relative that was also sequenced [diagnostic yield in family-based designs 37.4% (201/538) versus 15.2% (195/1287) in singletons; $P < 0.001$, Fisher's exact test].

We identified similar overall diagnostic rates at the time of each intake cut-off of our programme [19.1% (135/708) for 2019,[7] 20.0% (220/1100) for 2021,[8] 21.7% (396/1825) for 2024; Fig. 2A and B], with increased yields for 2019 and 2021 when implementing reanalysis strategies [1.0% (7/708) and 2.5% (28/1100) increase, respectively; Supplementary Table 2]. In the total cohort, we carried out reanalyses in an *ad hoc* manner, usually in the context of gene discovery,[30] variant classification updates[45] or filtering modifications for individual genotypes.[67] Dual diagnoses were achieved for nine patients (9/1825, 0.5%; Supplementary Table 2), indicating the presence of

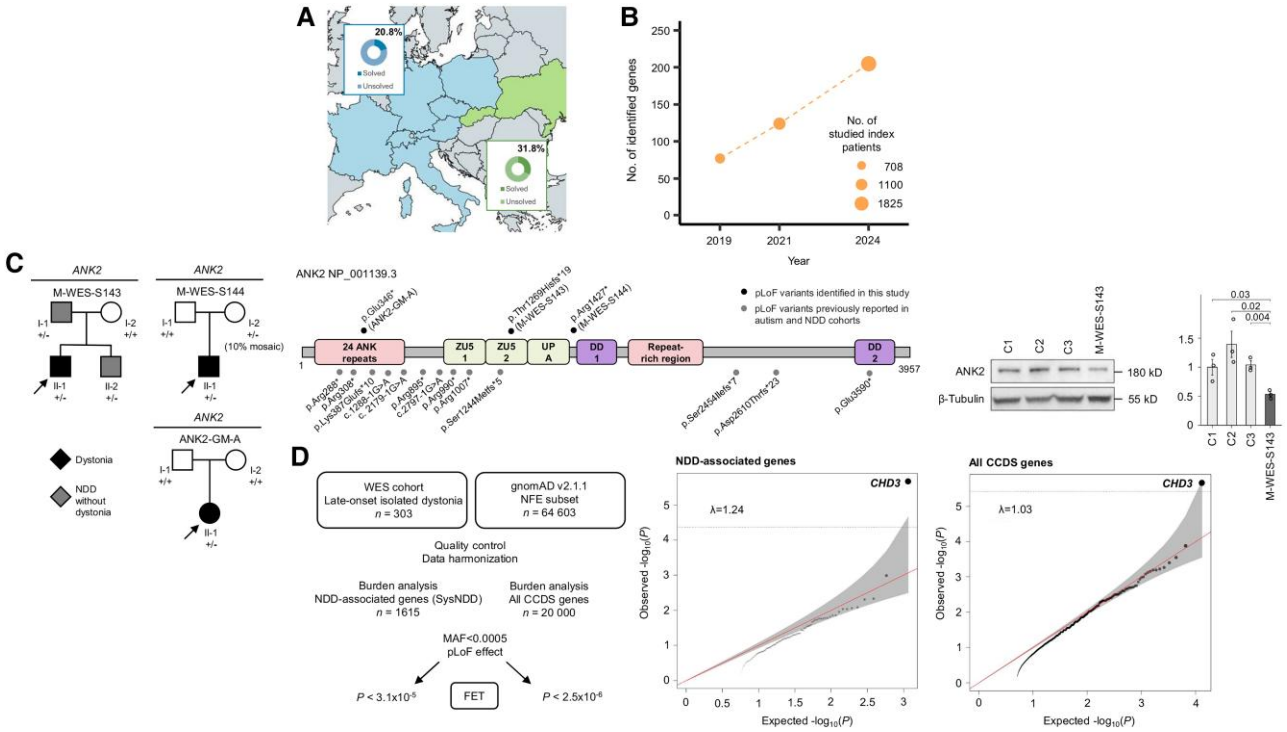

Figure 2 Linearly growing number of dystonia-associated genes in the WES cohort. (**A**) Recruitment sites in the study with inclusion of normally underrepresented patient groups and the associated overall diagnostic rates achieved by WES. Geographical areas with underserved dystonia populations and new recruitment foci are highlighted in green. (**B**) Cumulative dystonia patient ascertainment for WES over time, with the number of index patients analysed at the time cut-offs of the study in 2019[7]–2021[8]–2024 represented by the size of each point; a detailed description of the recruitment process is provided elsewhere.[7] The number of identified disease genes increased with increasing cohort size with no signs of plateauing. The diagnostic yield was relatively stable at ~19%–22%. Disease genes identified by WES in the entire cohort over a decade are as follows [alphabetical order; 141/205 (68.8%) found in a single family only; for details, see Supplementary Table 2]: *AARS1*, *ACTB*, *ADAR*, *ADCY5*, *AFG3L2*, *ALS2*, *ANK2*, *ANO3*, *AOPEP*, *ARHGEF9*, *ARSA*, *ASXL3*, *ATL1*, *ATM*, *ATP1A3*, *ATP2B2*, *ATP5F1A*, *ATP5F1B*, *ATP5MC3*, *ATP7B*, *ATP8A2*, *AUTS2*, *BCL11B*, *BRAF*, *BRPF1*, *C19orf12*, *CACNA1A*, *CACNA1E*, *CAMK4*, *CAMTA1*, *CASK*, *CD40LG*, *CHD3*, *CHD4*, *CHD8*, *CNTNAP1*, *COQ8A*, *CP*, *CSDE1*, *CTNNB1*, *CUL3*, *CUX1*, *CWF19L1*, *DCAF17*, *DDC*, *DHCR24*, *DHDDS*, *DLG4*, *DLL1*, *DNAJC6*, *DNM1L*, *DNMT1*, *EBF3*, *ECHS1*, *EEF1A2*, *EFTUD2*, *EIF2AK2*, *EIF4A2*, *ERCC4*, *ERCC8*, *FA2H*, *FBXO31*, *FGF14*, *FITM2*, *FOXG1*, *FOXP2*, *FRMD5*, *FRYL*, *FTL*, *GABBR2*, *GABRA1*, *GAD1*, *GCH1*, *GJA1*, *GJC2*, *GNAL*, *GNAO1*, *GNB1*, *GRIA2*, *GRIA3*, *GRID2*, *GRIN1*, *GRIN2A*, *HECW2*, *HEXA*, *HIBCH*, *IFIH1*, *IMPDH2*, *INTS11*, *IRF2BPL*, *KCNA2*, *KCNB1*, *KCNJ10*, *KCNMA1*, *KCTD17*, *KIF1A*, *KIF5A*, *KMT2B*, *LIG4*, *LRRK2*, *MAG*, *MATR3*, *MECP2*, *MECR*, *MED23*, *MICU1*, *MMAA*, *MORC2*, *MRE11*, *MSL3*, *NAA15*, *NARS2*, *NAV3*, *NEFL*, *NFIX*, *NGLY1*, *NKX2-1*, *NPC1*, *NR4A2*, *NUP54*, *OPA1*, *PAK1*, *PANK2*, *PARK7*, *PCDH12*, *PDE10A*, *PDHA1*, *PINK1*, *PLA2G6*, *PNKD*, *PNPLA6*, *POGZ*, *POLG*, *POLR1A*, *POLR3A*, *PPP2R5D*, *PPT1*, *PRKCG*, *PRKN*, *PRRT2*, *PSEN1*, *PTS*, *PURA*, *RALA*, *RARB*, *RERE*, *RHOBTB2*, *SATB1*, *SCN2A*, *SCO2*, *SCP2*, *SERAC1*, *SETX*, *SGCE*, *SHANK3*, *SHQ1*, *SLC16A2*, *SLC19A3*, *SLC20A2*, *SLC2A1*, *SLC6A1*, *SLC6A3*, *SLC9A6*, *SNAP25*, *SNX14*, *SON*, *SOX2*, *SOX6*, *SPAST*, *SPG11*, *SPG7*, *SPR*, *SPTBN1*, *SRRM2*, *SUCLG1*, *SUOX*, *SYNE1*, *TBC1D24*, *TBCD*, *TBX1*, *TCF20*, *TECPR2*, *TFE3*, *TH*, *THAP1*, *TMEM240*, *TOR1A*, *TTPA*, *TUBB4A*, *UBE3A*, *UBTF*, *VLDLR*, *VPS16*, *WAC*, *WARS2*, *WASHC5*, *WDR45*, *WDR73*, *WFS1*, *YY1*, *ZC4H2*, *ZEB2*, *ZMYND11*, *ZNF142*, *ZNF335*. (**C**) Pedigrees for three families (black fill, individual with dystonia; grey fill, individual with neurodevelopmental phenotype without dystonia; index patients indicated with arrows) with *ANK2* heterozygous predicted loss-of-function (pLoF) variants, and the positions of the variants mapped to the ANK2 protein sequence; *ANK2* pLoF mutations previously reported in autism[63] and other NDDs[64] are shown in the protein schematic for comparison (*bottom*). Patient ANK2-GM-A was identified via GeneMatcher.[65] The dystonia-related *ANK2* variants p.Glu346* and p.Thr1269Hisfs*19 were absent from gnomAD v4.1.0; p.Arg1427* was present in a single gnomAD-v.4.1.0 subject, as seen for an increasing number of NDD-causing variants associated with variable expressivity[41]; p.Arg1427* has also been identified in independent clinically affected individuals (listed as 'likely pathogenic' in ClinVar, ID: 3338732). Immunoblotting performed on fibroblasts from one WES-cohort individual (M-WES-S143) and three controls (C1–C3) showed significantly reduced ANK2 expression in the patient, compatible with the described haploinsufficiency mechanism of the disorder.[64] Blots are representative of three biological replicates (see also Supplementary Fig. 1); in bar plots for quantification, results are shown as mean ± standard deviation represented by error bars (statistical significance determined by Student's *t*-test). +/− = monoallelic variant carrier; +/+ = homozygous reference allele; *ANK2*, NM_001148.6. (**D**) Burden testing to demonstrate significant enrichment of heterozygous *CHD3* pLoF variants in adult-onset isolated dystonia. Differences in carrier rates of rare (MAF < 0.0005) pLoF variants (defined as nonsense, frameshift and splice-site alterations) between all individuals with late adulthood-onset[1] isolated dystonia in the WES cohort ($n = 303$) and controls from gnomAD-v.2.1.1 [non-Finish European (NFE) subset, $n = 64\,603$][40] were determined according to established methods (Test Rare vAriants with Public Data/TRAPD approach).[36,37] A flow chart of the analysis strategy is shown. Quantile-quantile plots for the study of NDD-associated genes ($n = 1615$; SysNDD database[66]) and all CCDS genes ($n = 20\,000$)[36] highlight a single significant signal for *CHD3* ($P < 3.1 \times 10^{-5}$ and $P < 2.5 \times 10^{-6}$, respectively, indicated with dashed horizontal lines; Fisher's exact test; genomic inflation factor λ is provided). CCDS = consensus coding sequence; gnomAD = Genome Aggregation Database; MAF = minor allele frequency; NDD = neurodevelopmental disorder; WES = whole-exome sequencing; WGS = whole-genome sequencing.

blended phenotypes related to the clinical spectra of different genes in large dystonia cohorts. Notably, there was higher diagnostic success in patients from populations with under recruitment, although the numbers were relatively small (49/154, 31.8%; Fig. 2A and Supplementary Table 2). Only 64 genes (64/205; 31.2%) had disease-related variants in two or more families (Supplementary Table 2), accounting for 65.7% of diagnosed index patients (260/396). All remaining diagnoses were unique, illustrating the immense diversity of underlying Mendelian causes in dystonia. When analysed by time intervals of the study (Fig. 2B), the number of

causally involved genes seemed to grow linearly, from 77 disease genes in the cohort in 2019 (708 families)[7] to 205 disease genes in 2024 (1825 families). Of these currently reported genes, 150 (73.2%) had established or proposed roles in brain development ('definitive' neurodevelopmental disorder-associated genes according to the SysNDD database[66]; Supplementary Table 2), consistent with our previous observations.[7] Each doubling of sample size for patients with dystonia and coexisting neurologic and/or systemic features corresponded to an ~90% increase in contributing neurodevelopmental loci.

While the rate of neurodevelopmental findings in patients ≥18 years of age was significantly lower than in the paediatric subgroup [10.0% (125/1245) versus 26.4% (153/580) diagnosed index patients; $P < 0.001$, FET], we encountered a number of unexpected diagnoses related to neurodevelopmental genes with no previous link to dystonia in cases from adult neurology practice. In two unrelated patients in young adulthood with dystonia, myoclonus and neuropsychiatric comorbidity, we detected truncating variants in *ANK2* (Fig. 2C and Supplementary Table 2), a gene in which pLoF mutations have originally been found in individuals with autism[63] and more recently in variable neurodevelopmental syndromes with epilepsy.[64] Reduced ANK2 expression, demonstrated in immunoblotting analysis, confirmed the LoF effect of the variant found in the patient with available fibroblasts (Fig. 2C and Supplementary Fig. 1). GeneMatcher-supported[65] identification of a patient with dystonia and epilepsy harbouring another *ANK2* pLoF alteration provided further evidence for an expanded gene–phenotype association (Fig. 2C). Moreover, we observed four mutually unrelated patients aged 48–64 years with isolated dystonia beginning in late adulthood[1] who had pLoF variants in *CHD3* (Supplementary Table 2), suggesting that disruption of this neurodevelopmental disorder gene[68] may have caused these subjects' phenotypes. To validate this, we performed gene-burden analyses of rare pLoF variants[36,37] across known neurodevelopmental disorder genes ($n = 1615$; SysNDD database,[66] accessed October 2023) and all coding genes[36] (Fig. 2D). When comparing pLoF variant counts in patients with late-onset (>40 years[1]) isolated dystonia in our cohort ($n = 303$) with the reference dataset, only *CHD3* but no other gene surpassed the multiple testing-corrected significance threshold in these tests ($P < 3.1 \times 10^{-5}$ for neurodevelopmental disorder-related genes, $P < 2.5 \times 10^{-6}$ for all genes; Fig. 2D).

Increase in diagnostic yield by whole-genome sequencing analysis

Second-line WGS enabled genetic diagnosis in 37 index patients (37/305, 12.1%; Table 1 and Supplementary Table 4). We identified 42 variations across 33 genes and 1 microduplication that we regarded with high confidence as causative (Table 1 and Supplementary Table 4). Of the 38 diagnoses (one patient had a dual diagnosis), 81.6% (31/38) were established on the basis of likely pathogenic/pathogenic[34,35] SNVs ($n = 12$), indels ($n = 4$), MT variants ($n = 4$) and CNVs/SVs ($n = 15$); six (6/38, 15.8%) were made on the basis of STR-alleles classified as disease causing ($n = 10$) and one on the basis of a pathogenic SNV plus an STR (Table 1, Supplementary Table 4 and Figs 3 and 4). Eighteen index patients (18/37, 48.6%) were found to have autosomal dominant disorders, 13 (13/37, 35.1%) autosomal recessive disorders, 2 (2/37, 5.4%) X-linked disorders and 4 (4/37, 10.8%) mitochondrial DNA-related disorders. We defined 7 of the 17 SNVs and indels (41.2%) as novel, because they had not been previously described in ClinVar[45] or the literature. The set of affected genes consisted of 28 known dystonia-associated genes (28/33, 84.8%) and 5 genes (5/33, 15.2%) that have been linked to neurologic diseases but not previously to dystonia.[13] Three index patients were considered to have ultra-rare conditions with fewer than 20 families reported to date (Supplementary Table 4). Review of the WGS findings identified that the solved patients had defied earlier diagnosis for a variety of reasons, including insufficient variant or disease gene evidence at the time of WES analysis,[13,14,45,46] lack of screening for variants outside the proximity of exons in clinical diagnostics,[71,72] technical inferiority of WES[73,74] and bioinformatic limitations of previously used pipelines[75,76] (Supplementary Table 4).

Eight diagnoses in seven index patients (7/37, 18.9%; Table 1 and Supplementary Table 4) were achieved by improved curation of coding variants, considering updated versions of ClinVar[45] and the latest gene–phenotype associations.[13,14] One example was *DNM1L*-related encephalopathy identified in two index patients with dystonia, developmental delay and epilepsy; the likely pathogenic heterozygous c.176C>T (p.Thr59Ile) missense variant, initially disregarded in WES analyses, was prioritized after a functionally validated alternative missense change at the same codon had been deposited in ClinVar (February 2022).[45] Known pathogenic intronic variants were observed in two index patients (2/37, 5.4%) who previously tested 'negative' by external WES; these were homozygous and compound heterozygous variants at +6 and +22 positions of donor sites in *POLR3A* (Table 1 and Supplementary Table 4). We uncovered reportable, previously unidentified variations in regions poorly captured by WES, including exonic sequences of *KMT2B* and *SHANK1*, resulting in different forms of neurodevelopmental dystonia, and we established a non-coding RNA gene-linked diagnosis by identification of compound heterozygous variants of *SNORD118* in a child with dystonia and leucoencephalopathy (Table 1, Supplementary Table 4 and Supplementary Figs 2 and 3). A mitochondrial genome analysis, previously unaccomplished on our WES platform, led to diagnoses in four additional index patients (4/37, 10.8%), associated with different levels of heteroplasmy (12%–97%) of pathogenic variants in *MT-ATP6*, *MT-ND3*, *MT-ND6* and *MT-TL1* (Table 1, Supplementary Table 4 and Supplementary Fig. 4).

We discovered disease-associated CNVs in 13 families (13/37, 35.1%), which were characterized with precise single-nucleotide breakpoint information (Table 1, Supplementary Table 4, Fig. 3 and Supplementary Figs 5 and 6). Eleven index patients (11/13, 84.6%) presented deletions (six heterozygous, one hemizygous, two homozygous, three compound heterozygous), ranging in size from 218 bp to 488.2 kb, and two carried heterozygous duplications (10.2 kb and 1.3 Mb) (Fig. 3A–E and Supplementary Figs 5 and 6). We reviewed the properties of these newly recognized CNVs: four were single-exon or partial-exon deletions (<5 kb; 4/14, 28.6%), explaining why these events were missed by previous screening;[76–78] similarly, overlapping compound heterozygous *PRKN* deletions, a deletion of the last two exons of *ATM*, as well as a *SGCE* 3-exon tandem duplication escaped WES-based detection in our bioinformatic pipeline; other CNVs were initially filtered out because of limited understanding of genotype-phenotype correlations, such as an *BCL11B* intra-exonic 218-bp deletion detected in two sisters with generalized dystonia (Fig. 3E); *BCL11B* has been associated with syndromic immunodeficiency, while only recent publications implied a link to dystonia-predominant phenotypes.[69,70] In three cases from the externally recruited subcohort (Fig. 1A), CNVs were not systematically captured from the WES data prior to WGS (Table 1 and Supplementary Table 4). Analysis of other SVs identified a *de novo* genomic rearrangement with breakpoints in coding and non-coding parts of the neurodevelopmental disease and

Table 1 Summary of genes with causative variants in 44 index patients according to WGS data analysis strategy

Number of solved index patients (total index patients analysed)	Identified disease genes (index patient study ID, zygosity of detected variant, variant category, variant size)
New variant or gene evidence (coding SNVs/indels)	
7 (305)	*ANO3* (G114, het, SNV, 1 bp), *ATP6V1A* (G162, het, SNV, 1 bp), *DNM1L* (G096, het, SNV, 1 bp), *DNM1L* (G227, het, SNV, 1 bp), *KMT5B* & *SRRM2* (dual diagnosis in G281, het, indel, 4 bp & het, SNV, 1 bp), *VPS16* (G125 het, indel, 1 bp), *XPA* (G191, hom, indel, 14 bp)
Known pathogenic intronic mutations	
2 (305)	*POLR3A* (G110, hom, SNV, 1 bp), *POLR3A* (G172, het/het, SNV/SNV, 1 bp/1 bp)
Coding variants not (sufficiently) covered by WES	
3 (305)	*KMT2B* (G043, het, indel, 7 bp), *SHANK1* (G113, het, SNV, 1 bp), *SNORD118* (G132, het/het, SNV/SNV, 1 bp/1 bp)
Mitochondrial DNA mutations	
4 (305)	*MT-ATP6* (G055, 96% mutational load, MT variant, 1 bp), *MT-ND3* (G085, 97% mutational load, MT variant, 1 bp), *MT-ND6* (G139, 25% mutational load, MT variant, 1 bp), *MT-TL1* (G254, 12% mutational load, MT variant, 1 bp)
CNVs/SVs	
14 (305)	*ASXL3* (G274, het, complex rearrangement, 352 bp duplication and 3.2 kb inversion and 2.6 kb deletion), *ATM* (G205, het/het, deletion CNV/SNV, 17.1 kb/1 bp), *BCL11B* (G226, het, deletion CNV, 218 bp), *CACNA1A* (G173, het, deletion CNV, 3.3 kb), *DLL1* (G120, het, deletion CNV, 488.2 kb), *PRKN* (G204, het/het, deletion CNV/deletion CNV, 286.1 kb/306.8 kb), *SGCE* (G222, het, duplication CNV, 10.2 kb), *SLC16A2* (G059, hem, deletion CNV, 3.1 kb), *SPG7* (G275, hom, deletion CNV, 714 bp), *THAP1* (G225, het, deletion CNV, 154.4 kb), *TIMM8A* (G201, het, deletion CNV, 25.4 kb), *TNRC6B* (G234, het, deletion CNV, 73.2 kb), *TTC19* (G174, hom, deletion CNV, 248 bp), 22q11.2 (G266, het, duplication CNV, ~1.3 Mb)
STR diagnoses	
7 (305)	*CSTB* (G236, hom, STR, CI: 27–51 CCCCGCCCCGCG units for allele 1 and 22–42 CCCCGCCCCGCG units for allele 2), *CSTB* (G293, hom, STR, CI: 29–70 CCCCGCCCCGCG units for allele 1 and 23–56 CCCCGCCCCGCG units for allele 2), *FXN* (G079, hom, STR, CI: 108–216 GAA units for allele 1 and 87–181 GAA units for allele 2), *FXN* (G299, hom, STR, CI: 107–228 GAA units for allele 1 and 86–189 GAA units for allele 2), *GLS* (G258, het/het, STR/SNV, CI: 115–195 GCA units/1 bp), *HTT* (G014, het, STR, 40 CAG units), *PABPN1* (G075, het, STR, 4 GCG and 3 GCA units)
Diagnoses enabled by integrative multi-omic analysis[a]	
6 (70 WGS-inconclusive index patients with available fibroblasts), plus 1 reanalysed case	*ATM* (G277, hom, MNV, 2 bp), *IRF2BPL* (G052, het, indel, 23 bp), *MECP2* (G161, het, indel, 47 bp), *SLC16A2* (G168, het, SNV, 1 bp), *SPG11* (G196, het/het, SNV/SNV, 1 bp/1 bp), *UFC1* (G245, het/het, indel/SNV, 12 bp/1 bp), *UFC1* (G269, het/het, indel/SNV, 1 bp/1 bp)

Additional details of the identified genes and variants, the sequencing designs, the associated phenotypes, the findings from fibroblast-based proteomics (and RNA sequencing) and reporting of barriers that were overcome by WGS or WGS plus proteomics are provided in Supplementary Table 4. bp = base pair(s); CI = confidence interval (ExpansionHunter); CNV = copy-number variant; hem = hemizygous; het = heterozygous; hom = homozygous; indel = short insertion/deletion (<50 bp); MNV = multinucleotide variation; MT variant = mitochondrial variant; SNV = single-nucleotide variant; STR = short tandem repeat; SV = structural variant; WES = whole-exome sequencing; WGS = whole-genome sequencing.

[a]Proteomics was performed to facilitate interpretation of variants of uncertain significance or variants in technically challenging regions and to guide variant prioritization in unresolved cases. Additional RNA sequencing was performed for patients with variants that were suspected to have a potential effect on splicing and/or transcript integrity.

dystonia-associated gene *ASXL3*,[79,80] which consisted of an inverted duplicated fragment and a deletion resulting in the disruption of two exons (Supplementary Fig. 7).

EH-based STR calling[31] revealed diagnoses for seven more families (7/37, 18.9%; Table 1, Supplementary Table 4, Fig. 4 and Supplementary Fig. 8). A polyglutamine expansion in *HTT* was found in a patient with dystonia and signal alterations of the basal ganglia who was clinically suspected of having a brain iron-accumulation disorder but not Huntington disease (Fig. 4A). In a family with multiple members affected by dystonia and muscular atrophy, the index patient and two affected siblings were identified with an expanded *PABPN1* polyalanine tract (Fig. 4B), underlying oculopharyngeal muscular dystrophy-1, a condition that may manifest with movement disorders, albeit rarely.[81] Bi-allelic dodecamer expansions in the 5′-untranslated region (UTR) of *CSTB* were called in two sisters with a similar phenotype involving myoclonic ataxia and dystonia (Fig. 4C); further assessment based on proteomics and transcriptomics showed significantly decreased expression of *CSTB* in both sisters' fibroblasts, confirming the diagnosis of Unverricht–Lundborg disease (ULD). Additionally, an unrelated index patient presented with dystonia-ataxia, myoclonus, and epilepsy and was identified to carry the expanded ULD-causing *CSTB* alleles (Table 1 and Supplementary Table 4). Another 5′-UTR repeat expansion was identified in *GLS*, *in trans* with a c.1197+2T>C splice-site SNV, in three siblings with a neurodevelopmental dystonia-ataxia syndrome and elevated plasma glutamine (1510 µmol/l in the index patient, reference range: 329–976 µmol/l); RNA-seq uncovered the presence of an abnormally extended exon-10 as a result of c.1197+2T>C, and the diagnosis of glutaminase deficiency was validated by biochemical studies: we demonstrated loss of GLS expression in immunoblotting on the index patient's fibroblasts (Fig. 4D and Supplementary Fig. 9), associated with markedly diminished GLS activity in patient cells

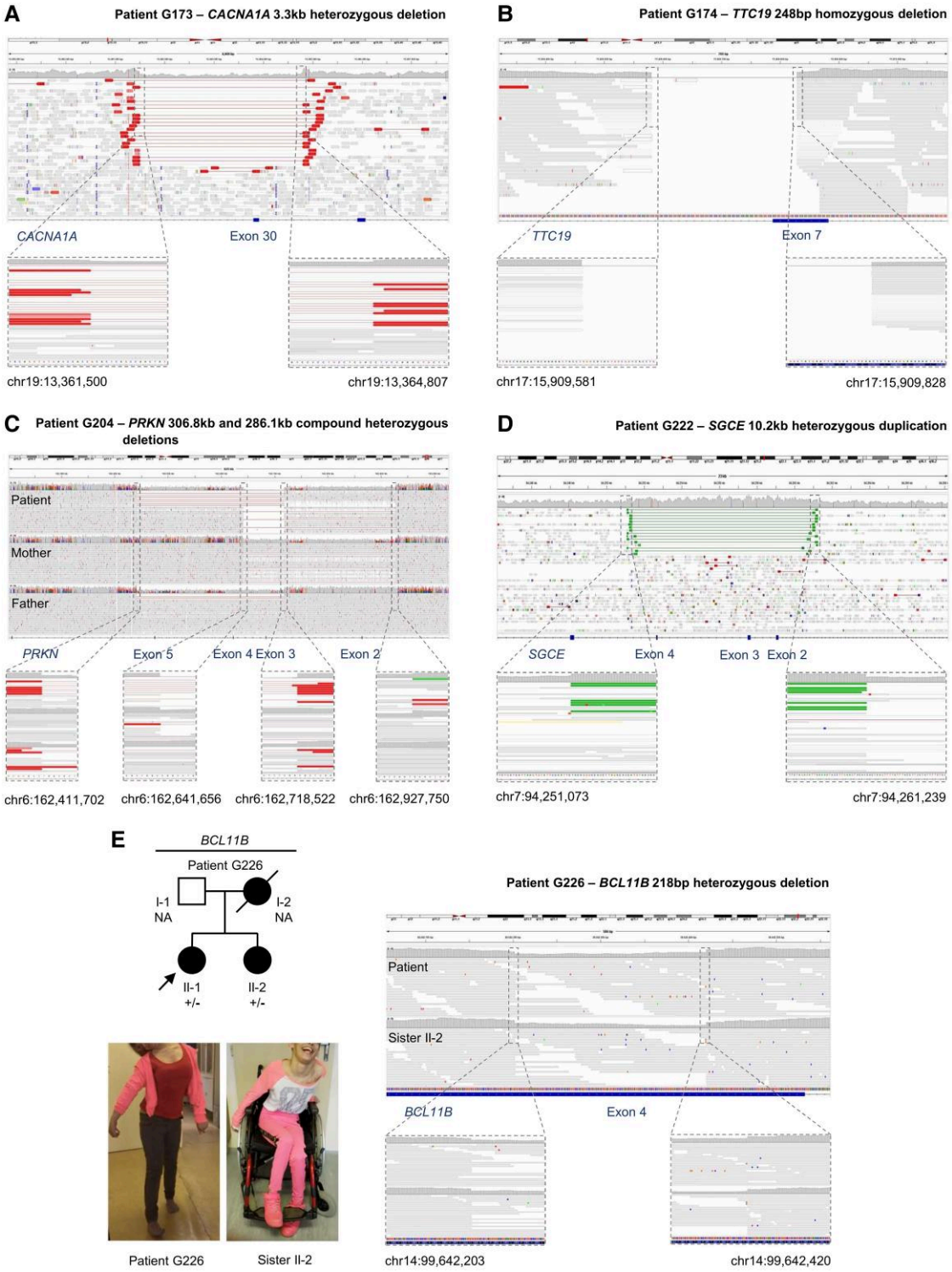

Figure 3 Genomic alignment of WGS data supporting CNVs reported in this study. Integrative Genomics Viewer[50] (IGV) screenshots are shown along with zoom-in panels depicting genomic breakpoints of CNVs. Uniform coverage allowed for resolution at the nucleotide level; genomic coordinates (hg19) of the breakpoint positions are provided. (**A**) A heterozygous single-exon deletion in *CACNA1A* (NM_000068.4: exon 30) in index patient G173 with dystonia and ataxia. (**B**) A homozygous deletion interrupting parts of *TTC19* exon-7 (NM_017775.4) in index patient G174 with dystonia, ataxia and intellectual disability. (**C**) Two overlapping heterozygous deletions in *PRKN* (NM_004562.3) in index patient G204 with dystonia-Parkinsonism. A 2-exon deletion (exons 2–3) was maternally inherited and a 3-exon deletion (exons 3–5) was paternally inherited. (**D**) A heterozygous intragenic tandem duplication in *SGCE* (NM_003919.3: exons 2–4) in index patient G222 with dystonia and myoclonus. (**E**) Family pedigree for index patient G226 (arrow) with dystonia and mild intellectual impairment; her sister (II-2) and deceased mother presented similar phenotypes. Representative clinical photographs of the two affected siblings illustrate generalized dystonic postures. A heterozygous intra-exonic frameshift deletion CNV was detected in exon 4 of *BCL11B* (NM_138576.4), the gene's hotspot for disease-causing truncating variants.[69,70] CNV = copy-number variant; NA = biological sample unavailable; WGS = whole-genome sequencing; +/− = monoallelic variant carrier.

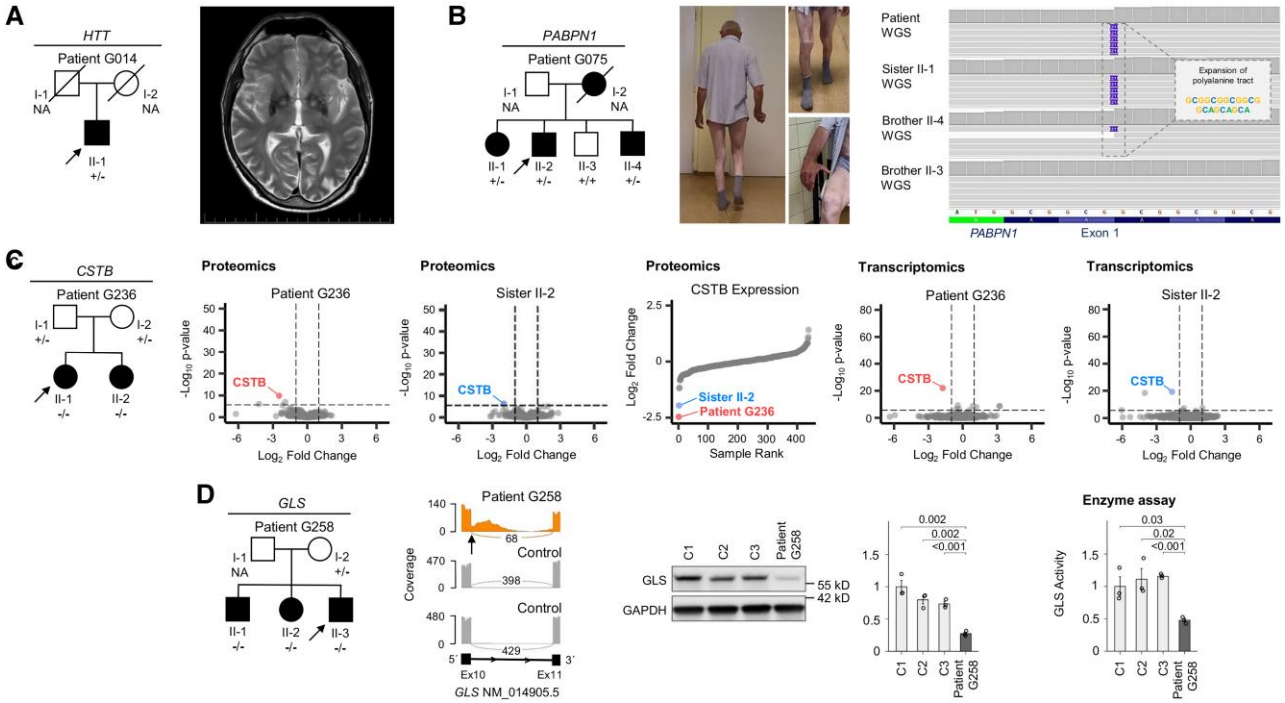

Figure 4 Four examples of repeat-expansion disorders diagnosed by WGS. (**A**) Family pedigree for index patient G014 (arrow) with dystonia, choreatic movements and T2 signal abnormalities in basal ganglia on cerebral MRI ('eye-of-the-tiger' aspect), suggestive of neurodegeneration with brain-iron accumulation disorder. ExpansionHunter[31] (EH) detected an expanded CAG allele with 40 repeat units in *HTT* (NM_001388492.1). (**B**) Family pedigree for index patient G075 (arrow) with dystonia, chorea, muscle weakness, dysphagia and cognitive decline; two siblings (II-1 and II-4) and the deceased mother had identical clinical disorders. Representative clinical photographs illustrate dystonic posturing while walking, as well as wasting of lower limbs and upper limbs in G075. EH-guided[31] inspection of WGS reads in Integrative Genomics Viewer[50] (IGV) revealed a 21-bp insertion, leading to an expansion of a polyalanine tract, in exon 1 of *PABPN1* (NM_004643.4) in the three affected siblings, but not in a fourth, healthy brother (II-3). (**C**) Family pedigree for index patient G236 (arrow) and her sister (II-2), both affected by epilepsy and a combined movement disorder with dystonia, ataxia and myoclonus. Proteomic and transcriptomic analyses were performed on fibroblasts[23,57] from G236 and II-2 to demonstrate a loss-of-function effect of EH-identified[31] bi-allelic CCCCGCCCCGCG expansions in the 5'-UTR of *CSTB* (NM_000100.4). Volcano plots of proteomics display significant underexpression of CSTB protein in both siblings. Rank plot comparing CSTB levels across all proteome samples[23,57] highlights the siblings as underexpression outliers; G236: CSTB, fold-change = 0.18, sample rank = 1; II-2: CSTB, fold-change = 0.26, sample rank = 2. Volcano plots of transcriptomics[55,57] display significant underexpression of *CSTB* mRNA in both siblings. In volcano plots, vertical lines represent $\log_2$fold-changes of −1 and 1, and horizontal lines indicate $P = 2.5 \times 10^{-6}$ (Bonferroni corrected *P*-value for 20 000 hypotheses corresponding to the number of theoretically identifiable gene-derived proteins/RNAs). The red and blue points represent the patient measurements. (**D**) Family pedigree for index patient G258 (arrow) and two similarly affected siblings with developmental delay and ataxic-dystonic movement disorders. In WGS data of all three siblings, EH screening[31] uncovered a 5'-UTR GCA expansion as a 'second hit' in *GLS* (NM_014905.5), in addition to an exon-10 c.1197+2T>C splice donor variant. A multimodal experimental approach was deployed to validate the diagnosis of 'global developmental delay, progressive ataxia and elevated glutamine' (MIM: 618 412) due to glutaminase (GLS) deficiency; first, plasma amino acids were analysed in G258 to reveal increased levels of the diagnostic marker[52] glutamine (1510 µmol/l, reference range: 329–976 µmol/l); second, fibroblast RNA-seq data were evaluated to identify a splicing abnormality resulting from c.1197+2T>C (black arrow): extension of exon-10 was observed in G258 but not in controls on Sashimi-plot visualization; third, immunoblotting was performed, showing drastic reduction of GLS expression in fibroblasts from G258 relative to three control lines (C1–C3; blots representative of three biological replicates, see also Supplementary Fig. 9; statistical significance in bar plots for quantification determined by Student's *t*-test); finally, diminished GLS activity in G258 was confirmed by enzymatic testing in patient and control cells (C1–C3; activities shown in relation to activity in C1; each data-point indicates a biological replicate; statistical significance determined by Student's *t*-test). NA = biological sample unavailable; RNA-seq = RNA-sequencing; UTR = untranslated region; WGS = whole-genome sequencing; +/− = monoallelic variant carrier; −/− = bi-allelic variant carrier; +/+ = homozygous reference allele.

(Fig. 4D). Finally, pathological intronic trinucleotide repeats in *FXN* were discovered in two index patients with adolescence-onset dystonia and ataxia (Table 1 and Supplementary Table 4).

High-throughput proteomics for variant functionalization and prioritization

In order to test proteomics as a potential adjunct to WGS diagnostics, we explored fibroblast proteomes of 80 patients (76 index patients and 4 affected relatives) and correlated findings with genomic results (Fig. 1A, Table 1 and Supplementary Table 4). On average, we detected ~8000 proteins per sample,[23] corresponding to ~51% of all OMIM-morbid genes.[23] Of all dystonia-associated genes according to OMIM[13] (accessed December 2023), 59% (352/598) were expressed in ≥75% of samples and 63% (375/598) in ≥50% of samples.

First, proteomics was instrumental in resolving low-quality variant calls and a variant of uncertain significance, yielding three new diagnoses (Table 1, Supplementary Table 4 and Fig. 5). In a child with dystonia and neuroregression, extended WGS filter strategies identified a *de novo* 23-bp frameshift duplication in *IRF2BPL*, removed from previous analyses as failing quality parameters (Fig. 5A); the variant was located in a region prone to alignment challenges and sequencing errors; the patient's proteomic data showed IRF2BPL downregulation (FC: 0.43) consistent with the known haploinsufficiency pathomechanism of *IRF2BPL*-associated neurodevelopmental disorder,[82] strongly supporting the variant as a true-positive diagnostic hit. In the case of a dystonic female patient with a differential

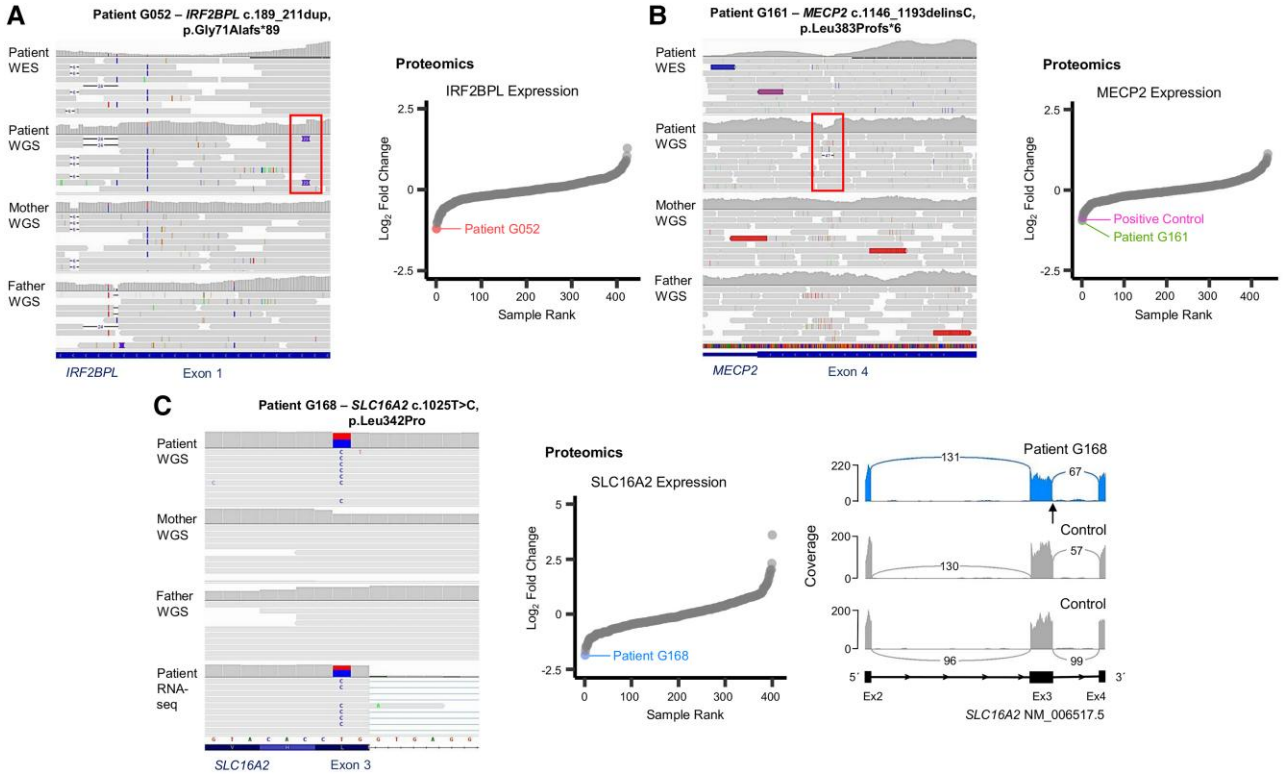

Figure 5 Diagnostic confidence of WGS results enhanced by proteomics. Integrative Genomics Viewer[50] (IGV) pileups of WGS-identified variants and rank plots for protein levels across all fibroblast proteome samples[23,57] are shown. The coloured points represent the patient measurements as indicated. (**A**) A *de novo* 23-bp duplication (red box) in a region with marked alignment complexity of *IRF2BPL* (NM_024496.4) in index patient G052. Note differences in read coverage of the target sequence between WES and WGS data. Diminished amounts of IRF2BPL protein in G052's proteome supported the diagnosis of 'neurodevelopmental disorder with regression, abnormal movements, loss of speech and seizures' (MIM: 618088), associated with heterozygous loss of IRF2BPL expression[82]; G052: IRF2BPL, fold-change = 0.43, sample rank = 1. (**B**) A poorly mapped *de novo* 47-bp indel (red box) in *MECP2* (NM_004992.4) in index patient G161. Note differences in read coverage of the target sequence between WES and WGS data. Underexpression of MECP2 protein confirmed the diagnosis of Rett syndrome. A sample with known pathogenic heterozygous loss-of-function *MECP2* variant[58] was included in the sample rank plot as a positive control; G161: MECP2, fold-change = 0.52, sample rank = 1; positive control: MECP2, fold-change = 0.55, sample rank = 2. (**C**) A *de novo* missense c.1025T>C (p.Leu342Pro) variant of uncertain significance in *SLC16A2* (NM_006517.5) identified in index patient G168. Low expression of SLC16A2 protein in G168's proteome was indicative of a loss-of-function effect of c.1025T>C, although the variant caused no observable splicing abnormality (see RNA-seq Sashimi plot, variant position indicated with black arrow); G168: SLC16A2, fold-change = 0.28, sample rank = 1. The diagnosis was further validated by demonstration of characteristic alterations on thyroid function testing[53] [elevated free T3 levels (7.2 pmol/l, reference range: 3.8–6 pmol/l) and elevated free T3/T4 ratio (0.82, reference range: <0.75)], adding G168 to the very short list of published female individuals with Allan–Herndon–Dudley syndrome.[83,84] RNA-seq = RNA sequencing; WES = whole-exome sequencing; WGS = whole-genome sequencing.

diagnosis of Rett syndrome, no suspicious variants were initially uncovered by WES/WGS; guided by the observation of relevant MECP2 reduction (FC: 0.52) in the patient's proteome, a poorly mapped (<20% of WGS reads) *MECP2 de novo* 47-bp frameshift indel, discarded by the analytical pipelines based on quality settings, was found after manual inspection in IGV (Fig. 5B). Another female patient, affected by dystonia and intellectual impairment, was identified to carry a *de novo* c.1025T>C (p.Leu342Pro) variant in the X-linked gene *SLC16A2* (Fig. 5C), classified as of uncertain significance because of the male preponderance in *SLC16A2*-related Allan–Herndon–Dudley syndrome[53,83-85] and the unknown consequence of a newly discovered missense substitution; proteomics indicated that the patient had the lowest abundance of SLC16A2 (FC: 0.28) among all study samples; in combination with characteristic thyroid hormone abnormalities[53] [elevated free T3 levels (7.2 pmol/l, reference range: 3.8–6 pmol/l) and elevated free T3/T4 ratio (0.82, reference range: <0.75)], this was considered convincing evidence to re-classify the variant as likely pathogenic and secured the diagnosis.

Second, untargeted outlier analysis of proteomic data pointed at four additional diagnostic findings that may not have otherwise been recognized (Table 1, Supplementary Table 4 and Fig. 6). We observed severely decreased amounts of SPG11 (FC: 0.01) in the proteome of a patient with dystonia and spasticity (Fig. 6A); reanalysis of WGS data revealed a *SPG11* nonsense variant plus an intronic c.3454-28A>G alteration, not retained by initial filtering; integrated RNA-seq demonstrated a splicing defect with exon skipping and intron retention for c.3454-28A>G, validating the diagnosis of spastic paraplegia-11. Similarly, decreased UFC1-protein levels (FC: 0.21) led to reprioritization of two corresponding variants, a *UFC1*4-amino acid deletion and an intronic c.255+17G>A variant, in the index patient of a family with three siblings affected by dystonia, developmental delay and spasticity (Fig. 6B); RNA-seq confirmed splice disruption with exon skipping, and segregation testing provided further evidence to support the diagnosis of *UFC1*-related neurodevelopmental disorder. A search for additional carriers of the *UFC1* intronic variant in our entire WGS-cohort datasets reidentified this mutation in compound heterozygosity with a frameshift variant in an unresolved patient with a spastic-dystonic neurodevelopmental syndrome (fibroblasts unavailable), indirectly establishing another diagnosis (Table 1 and Supplementary Table 4). Lastly, a proteome-wide search for expression changes of dystonia-associated proteins identified low levels of ATM (FC: 0.31) in an isolated dystonia-affected patient with

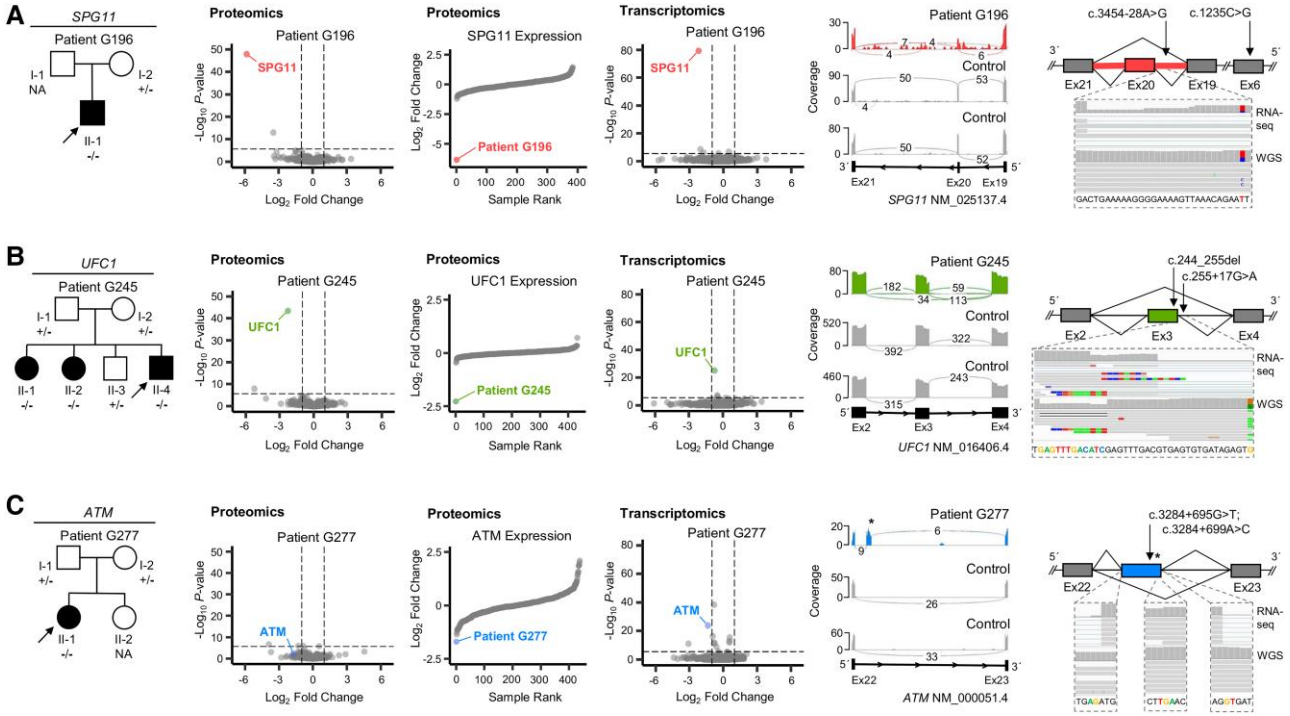

Figure 6 Proteomics-guided variant prioritization in WGS data. Unbiased outlier analysis of proteomic data[23] from index patient-derived fibroblast lines was performed and downregulated proteins were selected for closer examination in the context of WGS results. Volcano plots of fibroblast proteomic and transcriptomic analyses as well as rank plots for protein levels across all proteome samples[23,57] are shown. In volcano plots, vertical lines represent $\log_2$fold-changes of −1 and 1, and horizontal lines indicate $P = 2.5 \times 10^{-6}$ (Bonferroni corrected *P*-value for 20 000 hypotheses corresponding to the number of theoretically identifiable gene-derived proteins/RNAs). The coloured points represent the patient measurements as indicated. Schematic depictions not drawn to genomic scale. (**A**) Family pedigree for index patient G196 (arrow) with dystonia and spastic paraparesis. Proteomics highlighted SPG11 as expression outlier, associated with a combination of an intronic c.3454-28A>G variant and a nonsense mutation in *SPG11* (NM_025137.4) in WGS data. Transcriptomics confirmed *SPG11* underexpression at the RNA level. In Sashimi plots illustrating splicing in G196 (red) plus two representative controls, skipping of exon-20 and intron retention were observed as a consequence of *SPG11* c.3454-28A>G. The schematic depicts the variant locations and identified splicing abnormality (*top*; *bottom*: normal splicing pattern), and the embedded IGV captures of RNA-seq and WGS data show the read pileups at the position of c.3454-28A>G. (**B**) Family pedigree for index patient G245 (arrow) and two similarly affected siblings with developmental delay, generalized dystonia and spasticity. Proteomics highlighted UFC1 as expression outlier, associated with a combination of an intronic c.255+17G>A variant and a 4-amino acid deletion in *UFC1* (NM_016406.4) in WGS data. Transcriptomics confirmed *UFC1* underexpression at the RNA level. In Sashimi plots illustrating splicing in G245 (green) plus two representative controls, skipping of exon-3 was observed as a consequence of the *UFC1* variation. The schematic depicts the variant locations and identified splicing abnormality (*top*; *bottom*: normal splicing pattern), and the embedded Integrative Genomics Viewer[50] (IGV) captures of RNA-seq and WGS data show the read pileups at the position of the variants. (**C**) Family pedigree for index patient G277 (arrow) with isolated dystonia and elevated AFP. Scrutiny of underexpressed products of dystonia-associated genes in proteomic data highlighted reduced levels of ATM, accompanied by a significant *ATM* expression deficit in transcriptomics. WGS data re-evaluation uncovered a non-coding homozygous c. 3284+695G>T; c.3284+699A>C variation in *ATM* intron 22 (NM_000051.3), associated with the splicing-in of a 216-bp pseudoexon containing a premature stop codon; see Sashimi plots for G277 (blue) plus two representative controls (stop codon indicated with an asterisk). The schematic depicts the variant location and identified pseudoexon inclusion (*top*; *bottom*: normal splicing pattern), and the embedded IGV captures of RNA-seq and WGS data show the read pileups at the cryptic splice sites and the premature stop codon (*TGA*). AFP = alpha-fetoprotein; NA = biological sample unavailable; RNA-seq = RNA sequencing; WGS = whole-genome sequencing; +/− = monoallelic variant carrier; −/− = bi-allelic variant carrier.

no clear candidates from WES/WGS (Fig. 6C); using this information, we evaluated transcriptomic data and identified an *ATM* pseudoexon inclusion event, responsible for significant underexpression of *ATM* at the RNA level; re-review of the patient's WGS-variant profile illuminated a homozygous deep-intronic multinucleotide variation (c.3284+695G>T; c.3284+699A>C) as the genomic alteration underlying ataxia-telangiectasia with elevated AFP levels in this patient.

Candidate gene discovery and *PRMT1* variants identified as a novel cause of neurodevelopmental dystonia

As part of our WGS workflow, we undertook candidate gene searches coupled to systematic data sharing, primarily through a matchmaking platform[65] and direct communication.[20,30,86] We were able to assign compelling candidacy to four *de novo* heterozygous variants and three bi-allelic homozygous variants predicted to cause premature termination (Supplementary Table 5). Of the seven genes affected, one (*SRRM4*) is the subject of a manuscript in preparation, whereas another five await 'matches' (*ETV1*, *MYO16*, *TMEFF1*, *TXLNG*) or have functional studies underway (*ADCY1*).

In the candidate *PRMT1*, unassociated with a Mendelian disorder in OMIM,[13] trio WGS revealed a *de novo* c.1033C>T (p.Arg345Trp) missense variant in a patient with dystonia and neurodevelopmental dysfunction who previously had unrevealing WES in external laboratories (Supplementary Table 5 and Fig. 7). Via multi-site collaboration, three additional individuals with overlapping neurodevelopmental presentations (one with manifesting dystonia) and *PRMT1 de novo* missense changes [c.745G>T (p.Ala249Ser), c.871G>A (p.Glu291Lys)] could be identified (Fig. 7A). For an outline of the clinical and

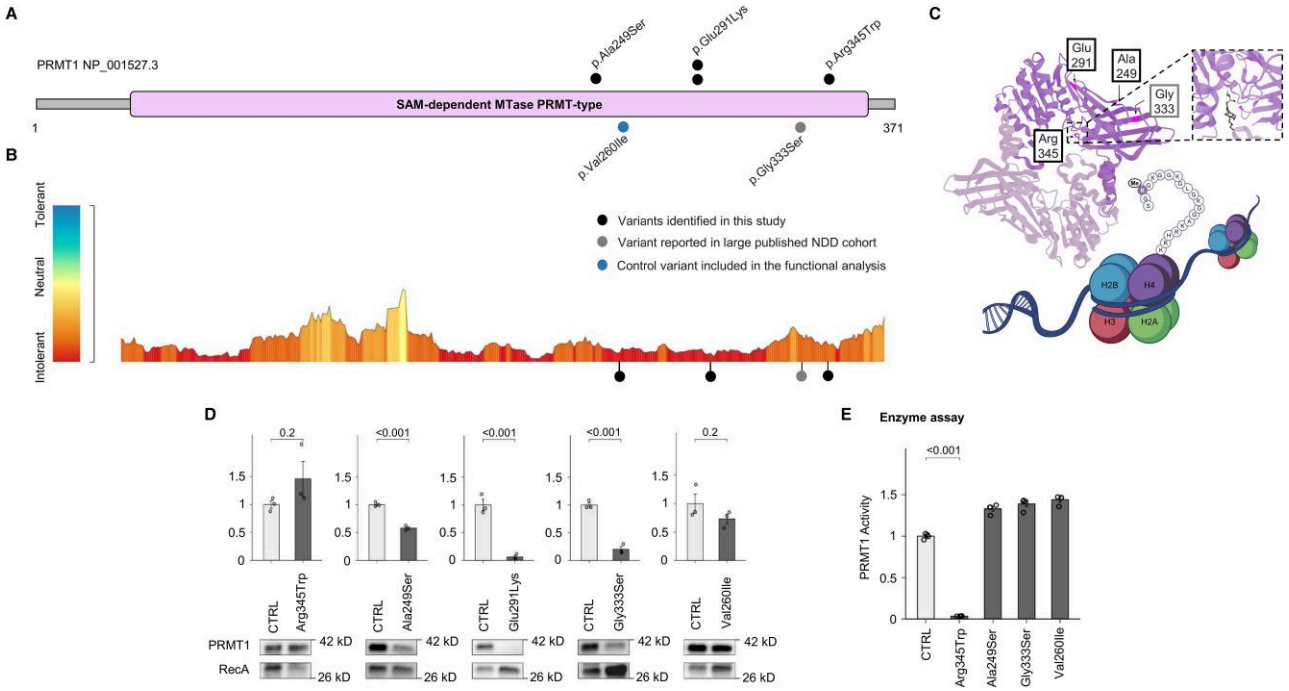

Figure 7 Discovery and functional characterization of *de novo* variants in *PRMT1*. (**A**) *PRMT1* variants (NM_001536.5) detected in four patients affected by neurodevelopmental phenotypes with or without dystonia mapped to the PRMT1 protein sequence (black, *top*). An additional *PRMT1* variant reported in a large study of *de novo* mutations in NDDs[60] is in grey (*bottom*); a gnomAD control variant[40] included in the *in vitro* assays is in blue (*bottom*). The canonical functional region is highlighted according to UniProt.[87] (**B**) Regional missense constraint over PRMT1, visualized via the MetaDome tolerance landscape server.[88] The relative positions of patient variants depicted in **A** are indicated on PRMT1 NP_938074.2, corresponding to the alternative transcript NM_198318.4 (MetaDome analysis not available for NM_001536.5). The variants were predicted to affect missense mutation-intolerant regions. (**C**) 3D dimer representation of PRMT1 (PDB: 6NT2) with mutated residues indicated; our variants (black boxes) were localized in the vicinity of the catalytic core and in the C-terminal β-barrel domain[89]; a magnified view of the region near residue Arg345 is depicted, illustrating the close spatial relationship between this mutated site and the binding pocket for PRMT1's cofactor S-adenosyl methionine (a synthetic inhibitor of this pocket is shown in stick representation).[90] The association with histones is illustrated, linking the identified disorder to the family of epigenetic-regulation defect syndromes, similarly to *KMT2B*-related dystonia.[91] (**D**) Representative immunoblots showing protein expression levels of normal control PRMT1 and mutant forms and bar plots for quantification. The here-identified variants p.Ala249Ser and p.Glu291Lys as well as the variant p.Gly333Ser from a published NDD cohort[60] reduced PRMT1 stability, indicating decreased functional PRMT1 levels in the carrier individuals. Stable expression was not significantly altered for PRMT1 proteins carrying the patient variant p.Arg345Trp or the gnomAD variant p.Val260Ile.[40] PRMT1 protein intensities were normalized to RecA. All analyses were performed in triplicate (each data-point indicates a biological replicate; see also Supplementary Fig. 10), and results are shown as mean ± standard deviation represented by error bars. Statistical analysis was performed by Student's *t*-test. (**E**) Enzymatic assay results for the PRMT1 missense substitutions, showing diminished activity compared with the normal control for the dystonia-associated variant p.Arg345Trp, but not for p.Ala249Ser, p.Gly333Ser,[60] and p.Val260Ile (gnomAD[40]). Means ± standard deviation are plotted, and statistical significance was determined by Student's *t*-test. gnomAD = Genome Aggregation Database; NDD = neurodevelopmental disorder.

molecular characteristics of each patient, see Supplementary Table 6. *PRMT1*, encoding an essential component of the epigenetic regulation machinery,[92] exhibits high population-scale intolerance against both missense and LoF alterations [gnomAD-v4.1.0: missense-z-score = 4.43, probability of being loss-of-function intolerant (pLI)-score = 1.0].[40] One of the identified variants [c.871G>A (p.Glu291Lys)] was recurrent in two patients. The three distinct variants all occurred at conserved residues in regions depleted in rare variation[88] (Fig. 7B) and none of them was found in control databases (gnomAD-v4.1.0, in-house collections). Bioinformatic predictions and three-dimensional-structural simulations were in favour of their pathogenicity (Fig. 7C and Supplementary Table 6). To evaluate directly whether the patient-specific *PRMT1* variants could be damaging to protein function, we investigated the expression levels and enzymatic activity of corresponding mutant constructs. One further *PRMT1 de novo* substitution published as part of a large neurodevelopmental disorder-cohort study (p.Gly333Ser)[60] and a gnomAD-listed variant (p.Val260Ile)[40] were also included in the experiments. As shown in Fig. 7D (see also Supplementary Fig. 10), stable expression of PRMT1 was substantially decreased for two samples bearing the variants p.Ala249Ser and p.Glu291Lys found in our patients in comparison to the normal control protein. The published mutation p.Gly333Ser[60] displayed similar behaviour, causing major changes in PRMT1's stability. Only one patient variant, p.Arg345Trp carried by the dystonia WGS-cohort case, had protein abundance similar to the normal control, but reached significantly reduced levels of enzyme function (Fig. 7E). The variants p.Ala249Ser and p.Gly333Ser had normal activity when the corresponding PRMT1 assay input was adjusted to normal control, suggesting that their residual expressed fractions were enzymatically intact (Fig. 7E), whereas p.Glu291Lys could not be tested in the enzymatic assay due to its severely compromised stability. In contrast, the variant from gnomAD did neither impair PRMT1 expression nor enzymatic ability. Together, these data indicated that the patient variants lowered the functional levels of PRMT1 or its activity, consistent with a LoF effect.

Discussion

We report on 10 years of genetic aetiology analyses in 1877 families with dystonia, integrating careful selection of eligible patients including extremely long cases in search of a diagnosis, WES, WGS,

adjunctive testing by other 'omics' modalities and data-sharing networks. The findings in aggregate showed that our concentration of expertise in investigating patients with an unmet diagnostic need provided a rich resource for elucidating the monogenic background of dystonia. WES performed by us achieved a cumulative diagnostic rate of 21.7% (396/1825), whereas WGS, dynamically combined with proteomic, transcriptomic and gene-discovery experiments, contributed to another 46 diagnoses (45/305 index patients, 14.8%). This investigation serves to highlight the potential of WGS and additional multi-omic analyses to improve insights into heterogeneous disease causes in patients with more severe forms of dystonia. Our experimental set-up allowed us to show a diagnostic uplift in such a subcohort characterized by higher levels of phenotypic complexity. Although many studies support the role of unbiased genome-wide approaches as a cornerstone for use in molecular diagnostics,[16,22,33] the benefit of WGS (and multi-omics) in unselected cohorts of WES-naive individuals with dystonia and patients with milder dystonic presentations remains to be explored. We appreciate that a lower aetiologic yield of advanced testing would be anticipated in the broader dystonia population.

Our present work had several strengths and points of difference compared with prior studies: (i) we established the utility of multifaceted molecular analysis strategies among families with a diverse range of dystonic syndromes on a scale far beyond any previously published cohort;[93] (ii) we showcase that elevation to WGS with simultaneous survey of multiple variant types represents an effective approach to diagnosis in dystonic patients for whom WES-based testing was unrevealing; (iii) we demonstrate that systematic implementation of proteomics is valuable for the interpretation of genomic findings in dystonia and can provide an additional increase in diagnostic yield of WGS [6/70 WGS-inconclusive index patients (8.6%), plus one reanalysed case; Table 1 and Supplementary Table 4]; and (iv) we expand the repertoire of epigenetic gene-related dystonia by online matchmaking and functional validation.

We sought to assess the added value from scrutiny of incremental numbers of WES datasets of patients from a wide spectrum of geographical origins. We found that the proportion of overlapping genes identified at different enrollment cut-offs in our study was considerably small [16.6% (34/205) for 2019[7] versus the period 2020–2024], suggesting that gene identification in dystonia will continue to meaningfully improve as cohort sizes grow. Our data reappraised the prevalence of dystonia-associated variants in neurodevelopmental genes,[94] discoverable across the whole age spectrum, including adults with broadened genotype-phenotype correlations for *ANK2* and *CHD3*. The observations amplified evidence predicting that hundreds, if not thousands of developmentally important genes may have the potential to be associated with dystonia.[13,94,95] Further, our experience with the large WES cohort implied that, in the future, it may be unlikely to find major high-penetrance genetic factors within coding regions shared by multiple undiagnosed patients, but rather a large residuum of contributing ultra-rare aetiologies will remain to be unravelled.

Follow-up investigations of the whole assayable genome allowed for diagnostic improvements that support the adoption of WGS after WES in unexplained cases with dystonia. WGS offered enhanced benefit by uncovering a wide set of previously unidentified or unprioritized CNVs/SVs (14 new diagnoses) in addition to WES-ignored SNVs and indels (13 diagnoses), and our WGS-based workflow enabled determination of the role of pathogenic MT variants in dystonia (four diagnoses), which has never been directly explored.[96] Moreover, we were surprised to identify seven index patients (7/305, 2.3%) with STR pathologies on WGS data, pointing to an underreported contribution of repeat-expansion disorders to dystonia causation.[97,98] Our results indicated that the boundary between dystonia and 'classical' expressions of STR-related diseases such as muscle weakness (e.g. in *PABPN1*-associated muscular dystrophy) or ataxia (e.g. in *CSTB*-related ULD) can be blurred, highlighting the difficulty of diagnosing rare multisystem neurologic disorders clinically.[99] Although our patients with STR-associated diagnoses had neurologic features characteristic of the respective conditions, dystonia was observed as a leading clinical abnormality in all these cases and was a main reason for inclusion in the WGS analysis programme. We performed an array of resource-intensive *post hoc* validation tests for reported STRs, stressing the need for further studies to reach consensus on STR screening and validity assessments in WGS experiments.[39]

In conditions that evaded detection with the sole application of WGS, gene-agnostic proteomics emerged as an important driver of diagnostic success in dystonia. First, identification of expression defects increased the diagnostic certainty of missense and indel variants, such as frameshift mutations in regions with low sequencing quality scores. We were unable to identify and/or validate the reported 23-bp duplication in *IRF2BPL* (Fig. 5A) through other analytic strategies including WES and Sanger testing, highlighting the advantage of our combined WGS-proteomics approach over more traditional techniques. In the case of the *SLC16A2* missense variant that we re-classified on the basis of proteomics (Fig. 5C), we note that a comparison of SLC16A2-protein expression levels between our studied patient and confirmed cases of Allan–Herndon–Dudley syndrome was not possible because positive-control samples were unavailable to us; nevertheless, our proteomic result guided the successful diagnostic outcome, as it was decisive for the initiation of ancillary evaluation of endocrinological parameters and complemented the clinico-molecular picture of SLC16A2 deficiency. Second, we benchmarked the proteomic assay as a robust tool to find dystonic patients with disease-associated intronic variants resulting in reduction of protein levels, overcoming pitfalls of conventional WGS analyses that do not decipher functional impact.[18,21,22] It is within reason to expect that proteomics will efficiently improve diagnostics of dystonia in parallel to other omics implementations such as RNA-seq.[20,29] Our study demonstrates how implementation of proteomics can expand the capabilities to perform efficient molecular analysis in a collection of prioritized patients. However, it should be emphasized that the diagnostic gains of such an approach in individuals with more common types of dystonia such as adult-onset isolated dystonia are unknown. In time, multi-omic studies are likely to become more broadly used as workflows will be improved and costs will continue to fall, which may enhance our knowledge about testing indications. Currently, one reasonable approach could be to consider the use of WGS plus multi-omics in patients with dystonic syndromes that are the most likely to be caused by single-gene disorders.

Rigorous candidate-gene prioritization informed by knowledge of molecular pathways[15,26] allowed us to pursue case-matching and characterization of a novel genetic aetiology for dystonia: the similar clinical manifestations of the carrier individuals and the functional results showing a loss-of-function mechanism provided conclusive support for the implication of *PRMT1* variants in a Mendelian phenotype with dystonia, reinforcing the connection between histone methyltransferase defects and neurodevelopmental disease.[13,60]

All positive results from WGS and WGS combined with proteomics informed genetic counselling. Many findings ended long journeys to diagnosis [diagnostic delay of >5 years in 72.7% (32/44) of index patients], with pronounced duration in individuals carrying disease-causing STRs and non-coding variation (average delay of

11 years). Moreover, the established diagnoses aided in making care decisions and suggested tailored management for 29 families (29/44, 65.9%; Supplementary Table 7). In light of prospects for mutation-specific therapies, we valued diagnoses related to splice-altering variants in *SPG11* and *ATM*, uniquely identified by WGS with integration of proteomics and RNA-seq (Fig. 6A and C): antisense oligonucleotides acting via splicing modulation are under development for these genes.[100] Specifically, the identification of deep intronic variants causing ataxia-telangiectasia may enable eligibility for a trial with targeted suppression of *ATM* cryptic exonization.[101]

There were weaknesses in our study: we acknowledge that our cohort still needs to include more patients from as-yet underrepresented populations.[102] Short-read WGS, as applied by us, has limitations in identifying certain variant types,[20] and the study was not primarily and, therefore, not perfectly designed to compare the diagnostic utility between WES and WGS. Fibroblast proteomics required invasive biopsies, and the approach had limited detection capacity for 37% of the dystonia-relevant gene products (products detected in <50% of samples).[23] The lack of expression of brain-specific disease genes and other confounding cellular events including tissue-dependent alternative splicing or alternative cleavage of transcripts represent obstacles in integrating fibroblast-based proteomics and RNA-seq into a streamlined diagnostic process for all patients. Future innovative approaches such as diagnostics of fibroblasts subjected to neural transdifferentiation[103] or clustered regularly interspaced short palindromic repeat (CRISPR)-mediated transcriptional activation[104] may help to further enhance our abilities to investigate the functional impact of genomic variants; however, expression-defect profiling alone would still be unable to assess some mutational consequences such as gain of protein function. Finally, our gene-sharing strategies were imperfect since some 'weaker' candidate loci and variants, e.g. dominant alleles inherited from asymptomatic parents, were not elected for submission to data-sharing services.

Overall, we demonstrate for a multinational cohort meticulously studied by sequencing and proteomics that a precision genetics approach delivered in a clinical–academic format can yield a sizable rate of accurate diagnoses for dystonia-affected families. Failure to make these diagnoses would have led to missed opportunities for optimized care, including lack of qualification for upcoming tailored interventions such as treatment with antisense technologies.[100,105] Our outcome illustrating the benefits of WGS and multi-omic tests for difficult-to-diagnose patients may serve as a roadmap for reconsideration of diagnostic algorithms, paving the way for innovative strategies to uncover the elusive genetic basis of dystonia.

Data availability

The data that support the findings of this study are available from the corresponding author, upon request.

Acknowledgements

We are deeply indebted to the patients and their families who have supported our research activities for years. The authors are grateful to the Deutsche Dystonie Gesellschaft (DDG), Dystonia Europe (DE) and the Dystonia Medical Research Foundation (DMRF). We thank Celestine Dutta, Katharina Eyring, Sandy Lösecke, Katharina Mayerhanser, Nicola Plathner, Elisabeth Stephan and Monika Zimmermann for excellent technical assistance. Moreover, we would like to acknowledge the assistance of the BayBioMS core facility of Technical University of Munich, Freising, Germany, and the Protein expression and purification facility, Helmholtz Munich, Munich, Germany. M.Z. is a member of the Medical and Scientific Advisory Council of the DMRF and a member of the Governance Council of the International Cerebral Palsy Genomics Consortium (ICPGC). We thank the Cell line and DNA bank of Genetic Movements Disorders and Mitochondrial Diseases of the Telethon Network of Genetic Biobanks (TNGB-P09). The proteomic data has been generated at BayBioMS, TU Munich, in collaboration with Christina Ludwig. We would like to thank the group of Julien Gagneur for their help in the development of the bioinformatic pipelines. We acknowledge the use of BioRender that was used to create parts of the thumbnail image.

Funding

This work was supported by the DFG Research Infrastructure NGS_CC (project #458949627) as part of the Next Generation Sequencing Competence Network (project 423957469); NGS analyses were carried out at the production site WGGC-Bonn; Michael Zech, Barbara Schormair and Juliane Winkelmann received the DFG grants ZE 1213/2-1, SCHO 1644/4-1, WI 1820/14-1 (#458949627) as part of the DFG Sequencing call Sequencing Costs in Projects. Our work was further supported by funding from the EJP RD (EJP RD Joint Transnational Call 2022) and the German Federal Ministry of Education and Research (BMBF, Bonn, Germany), awarded to the project PreDYT (PREdictive biomarkers in DYsTonia, 01GM2302) and GENOMIT. This research was supported by a '*Schlüsselprojekt*' grant from the Else Kröner-Fresenius-Stiftung (2022_EKSE.185). In addition, this study (M.Z.) has received funding from the Federal Ministry of Education and Research (BMBF) and the Free State of Bavaria under the Excellence Strategy of the Federal Government and the Länder, as well as by the Technical University of Munich—Institute for Advanced Study. Part of this research was made possible through access to the data (Technical University of Munich, Munich, Germany) generated by the Bavarian Genomes Network. The work was also supported by the National Institute for Neurological Research, Czech Republic, Programme EXCELES, ID Project No. LX22NPO5107, funded by the European Union—Next Generation EU and also by the Charles University: Cooperatio Program in Neuroscience and by the Agency for Health Research of the Czech Republic (grant AZV: NW24-04-00067). Moreover, financial support by the Slovak Grant and Development Agency under contract APVV-22-0279, by the EU Renewal and Resilience Plan 'Large projects for excellent researchers' under grant No. 09I03-03-V03-00007, and by the Slovak Scientific Grant Agency under contract no. VEGA 1/0712/22 to M.S., J.B., D.H. and M.O. is gratefully acknowledged. B.G. acknowledges grant support from Fondazione Mariani (CM23). Several authors are members of the European Reference Network for Rare Neurological Diseases (ERN-RND). E.I. acknowledges support from the European Joint Programme on Rare Diseases (EJP-RD WP17 research mobility fellowship). L.M.R. received grant support from the Italian Ministry of Health (GR-2009-1594645). A.D.F. acknowledges grant support by the EJP RD (EJP RD Joint Transnational Call 2022) and 'Fondazione Regionale per la Ricerca Biomedica' awarded to the projectPreDYT (PREdictive biomarkers in DYsTonia, grant agreement number: 825575). H.P. acknowledges grant support from the German Federal Ministry of Education and Research (BMBF, Bonn, Germany) awarded to the German Network for Mitochondrial Disorders (mitoNET, 01GM1906A).

Competing interests

The authors report no competing interests.

Supplementary material

Supplementary material is available at *Brain* online.

References

1. Albanese A, Bhatia K, Bressman SB, et al. Phenomenology and classification of dystonia: A consensus update. *Mov Disord*. 2013;28:863-873.
2. Grutz K, Klein C. Dystonia updates: Definition, nomenclature, clinical classification, and etiology. *J Neural Transm (Vienna)*. 2021;128:395-404.
3. van Egmond ME, Kuiper A, Eggink H, et al. Dystonia in children and adolescents: A systematic review and a new diagnostic algorithm. *J Neurol Neurosurg Psychiatry*. 2015;86:774-781.
4. van Egmond ME, Lagrand TJ, Lizaitiene G, Smit M, Tijssen MAJ. A novel diagnostic approach for patients with adult-onset dystonia. *J Neurol Neurosurg Psychiatry*. 2022;93:1039-1048.
5. Pozojevic J, Beetz C, Westenberger A. The importance of genetic testing for dystonia patients and translational research. *J Neural Transm (Vienna)*. 2021;128:473-481.
6. Sarva H, Rodriguez-Porcel F, Rivera F, et al. The role of genetics in the treatment of dystonia with deep brain stimulation: Systematic review and meta-analysis. *J Neurol Sci*. 2024;459:122970.
7. Zech M, Jech R, Boesch S, et al. Monogenic variants in dystonia: An exome-wide sequencing study. *Lancet Neurol*. 2020;19:908-918.
8. Dzinovic I, Boesch S, Skorvanek M, et al. Genetic overlap between dystonia and other neurologic disorders: A study of 1,100 exomes. *Parkinsonism Relat Disord*. 2022;102:1-6.
9. Ahn JH, Kim AR, Park WY, Cho JW, Park J, Youn J. Whole exome sequencing and clinical investigation of young onset dystonia: What can we learn? *Parkinsonism Relat Disord*. 2023;115:105814.
10. Dhar D, Holla VV, Kumari R, et al. Clinical and genetic profile of patients with dystonia: An experience from a tertiary neurology center from India. *Parkinsonism Relat Disord*. 2024;120:105986.
11. Atasu B, Simon-Sanchez J, Hanagasi H, et al. Dissecting genetic architecture of rare dystonia: Genetic, molecular and clinical insights. *J Med Genet*. 2024;61:443-451.
12. Thomsen M, Marth K, Loens S, et al. Large-scale screening: Phenotypic and mutational spectrum in isolated and combined dystonia genes. *Mov Disord*. 2024;39:526-538.
13. Hamosh A, Scott AF, Amberger JS, Bocchini CA, McKusick VA. Online Mendelian inheritance in man (OMIM), a knowledgebase of human genes and genetic disorders. *Nucleic Acids Res*. 2005;33(Database issue):D514-D517.
14. Stark Z, Foulger RE, Williams E, et al. Scaling national and international improvement in virtual gene panel curation via a collaborative approach to discordance resolution. *Am J Hum Genet*. 2021;108:1551-1557.
15. Thomsen M, Lange LM, Zech M, Lohmann K. Genetics and pathogenesis of dystonia. *Annu Rev Pathol*. 2024;19:99-131.
16. Wright CF, Campbell P, Eberhardt RY, et al. Genomic diagnosis of rare pediatric disease in the United Kingdom and Ireland. *N Engl J Med*. 2023;388:1559-1571.
17. Zech M, Jech R, Boesch S, et al. Scoring algorithm-based genomic testing in dystonia: A prospective validation study. *Mov Disord*. 2021;36:1959-1964.
18. 100,000 Genomes Project Pilot Investigators; Smedley D, Smith KR, et al. 100,000 genomes pilot on rare-disease diagnosis in health care—Preliminary report. *N Engl J Med*. 2021;385:1868-1880.
19. Wojcik MH, Reuter CM, Marwaha S, et al. Beyond the exome: What's next in diagnostic testing for Mendelian conditions. *Am J Hum Genet*. 2023;110:1229-1248.
20. Zech M, Winkelmann J. Next-generation sequencing and bioinformatics in rare movement disorders. *Nat Rev Neurol*. 2024;20:114-126.
21. Kumar KR, Davis RL, Tchan MC, et al. Whole genome sequencing for the genetic diagnosis of heterogenous dystonia phenotypes. *Parkinsonism Relat Disord*. 2019;69:111-118.
22. Bertoli-Avella AM, Beetz C, Ameziane N, et al. Successful application of genome sequencing in a diagnostic setting: 1007 index cases from a clinically heterogeneous cohort. *Eur J Hum Genet*. 2021;29:141-153.
23. Kopajtich R, Smirnov D, Stenton SL, et al. Integration of proteomics with genomics and transcriptomics increases the diagnostic rate of Mendelian disorders. *bioRxiv*. [Preprint] https://doi.org/10.1101/2021.03.09.21253187
24. Invernizzi F, Legati A, Nasca A, et al. Myopathic mitochondrial DNA depletion syndrome associated with biallelic variants in LIG3. *Brain*. 2021;144:e74.
25. Lunke S, Bouffler SE, Patel CV, et al. Integrated multi-omics for rapid rare disease diagnosis on a national scale. *Nat Med*. 2023;29:1681-1691.
26. Di Fonzo A, Jinnah HA, Zech M. Dystonia genes and their biological pathways. *Int Rev Neurobiol*. 2023;169:61-103.
27. Braconi D, Bernardini G, Spiga O, Santucci A. Leveraging proteomics in orphan disease research: Pitfalls and potential. *Expert Rev Proteomics*. 2021;18:315-327.
28. Smirnov D, Konstantinovskiy N, Prokisch H. Integrative omics approaches to advance rare disease diagnostics. *J Inherit Metab Dis*. 2023;46:824-838.
29. Prokisch H, Zech M. Proteomic profiling in dystonia: The next frontier for pathophysiology research and biomarker exploration. *Mov Disord*. 2024;39:1478-1479.
30. Harrer P, Skorvanek M, Kittke V, et al. Dystonia linked to EIF4A2 haploinsufficiency: A disorder of protein translation dysfunction. *Mov Disord*. 2023;38:1914-1924.
31. Dolzhenko E, Deshpande V, Schlesinger F, et al. ExpansionHunter: A sequence-graph-based tool to analyze variation in short tandem repeat regions. *Bioinformatics*. 2019;35:4754-4756.
32. Lim SY, Tan AH, Ahmad-Annuar A, et al. Uncovering the genetic basis of Parkinson's disease globally: From discoveries to the clinic. *Lancet Neurol*. 2024;23:1267-1280.
33. Stranneheim H, Lagerstedt-Robinson K, Magnusson M, et al. Integration of whole genome sequencing into a healthcare setting: High diagnostic rates across multiple clinical entities in 3219 rare disease patients. *Genome Med*. 2021;13:40.
34. 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.
35. Riggs ER, Andersen EF, Cherry AM, et al. Technical standards for the interpretation and reporting of constitutional copy-number variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics (ACMG) and the Clinical Genome Resource (ClinGen). *Genet Med*. 2020;22:245-257.
36. Steel D, Zech M, Zhao C, et al. Loss-of-function variants in HOPS complex genes VPS16 and VPS41 cause early onset dystonia associated with lysosomal abnormalities. *Ann Neurol*. 2020;88:867-877.
37. Guo MH, Plummer L, Chan YM, Hirschhorn JN, Lippincott MF. Burden testing of rare variants identified through exome sequencing via publicly available control data. *Am J Hum Genet*. 2018;103:522-534.

38. Coutelier M, Holtgrewe M, Jager M, et al. Combining callers improves the detection of copy number variants from whole-genome sequencing. *Eur J Hum Genet*. 2022;30:178-186.
39. Ibanez K, Polke J, Hagelstrom RT, et al. Whole genome sequencing for the diagnosis of neurological repeat expansion disorders in the UK: A retrospective diagnostic accuracy and prospective clinical validation study. *Lancet Neurol*. 2022;21:234-245.
40. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. 2020;581:434-443.
41. Indelicato E, Eberl A, Boesch S, et al. Genome aggregation database version 4-allele frequency changes and impact on variant interpretation in dystonia. *Mov Disord*. 2025;40:357-362.
42. Lappalainen I, Lopez J, Skipper L, et al. Dbvar and DGVa: Public archives for genomic structural variation. *Nucleic Acids Res*. 2013;41(Database issue):D936-D941.
43. MacDonald JR, Ziman R, Yuen RK, Feuk L, Scherer SW. The database of genomic variants: A curated collection of structural variation in the human genome. *Nucleic Acids Res*. 2014;42(Database issue):D986-D992.
44. Brandon MC, Lott MT, Nguyen KC, et al. MITOMAP: A human mitochondrial genome database–2004 update. *Nucleic Acids Res*. 2005;33(Database issue):D611-D613.
45. Landrum MJ, Lee JM, Benson M, et al. ClinVar: Public archive of interpretations of clinically relevant variants. *Nucleic Acids Res*. 2016;44(D1):D862-D868.
46. Stenson PD, Mort M, Ball EV, et al. The human gene mutation database: Towards a comprehensive repository of inherited mutation data for medical research, genetic diagnosis and next-generation sequencing studies. *Hum Genet*. 2017;136:665-677.
47. Chintalaphani SR, Pineda SS, Deveson IW, Kumar KR. An update on the neurological short tandem repeat expansion disorders and the emergence of long-read sequencing diagnostics. *Acta Neuropathol Commun*. 2021;9:98.
48. Ellingford JM, Ahn JW, Bagnall RD, et al. Recommendations for clinical interpretation of variants found in non-coding regions of the genome. *Genome Med*. 2022;14:73.
49. Bamshad MJ, Nickerson DA, Chong JX. Mendelian gene discovery: Fast and furious with no end in sight. *Am J Hum Genet*. 2019;105:448-455.
50. Robinson JT, Thorvaldsdottir H, Winckler W, et al. Integrative genomics viewer. *Nat Biotechnol*. 2011;29:24-26.
51. Waldmann TA, McIntire KR. Serum-alpha-fetoprotein levels in patients with ataxia-telangiectasia. *Lancet*. 1972;2:1112-1115.
52. van Kuilenburg ABP, Tarailo-Graovac M, Richmond PA, et al. Glutaminase deficiency caused by short tandem repeat expansion in GLS. *N Engl J Med*. 2019;380:1433-1441.
53. Remerand G, Boespflug-Tanguy O, Tonduti D, et al. Expanding the phenotypic spectrum of Allan-Herndon-Dudley syndrome in patients with SLC16A2 mutations. *Dev Med Child Neurol*. 2019;61:1439-1447.
54. Zech M, Kopajtich R, Steinbrucker K, et al. Variants in mitochondrial ATP synthase cause variable neurologic phenotypes. *Ann Neurol*. 2022;91:225-237.
55. Yepez VA, Gusic M, Kopajtich R, et al. Clinical implementation of RNA sequencing for Mendelian disease diagnostics. *Genome Med*. 2022;14:38.
56. Brechtmann F, Mertes C, Matuseviciute A, et al. OUTRIDER: A statistical method for detecting aberrantly expressed genes in RNA sequencing data. *Am J Hum Genet*. 2018;103:907-917.
57. Brugger M, Lauri A, Zhen Y, et al. Bi-allelic variants in SNF8 cause a disease spectrum ranging from severe developmental and epileptic encephalopathy to syndromic optic atrophy. *Am J Hum Genet*. 2024;111:594-613.
58. Pascual-Alonso A, Xiol C, Smirnov D, Kopajtich R, Prokisch H, Armstrong J. Identification of molecular signatures and pathways involved in Rett syndrome using a multi-omics approach. *Hum Genomics*. 2023;17:85.
59. Tyanova S, Temu T, Cox J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc*. 2016;11:2301-2319.
60. Kaplanis J, Samocha KE, Wiel L, et al. Evidence for 28 genetic disorders discovered by combining healthcare and research data. *Nature*. 2020;586:757-762.
61. Dummler A, Lawrence AM, de Marco A. Simplified screening for the detection of soluble fusion constructs expressed in *E. coli* using a modular set of vectors. *Microb Cell Fact*. 2005;4:34.
62. Zanuttigh E, Derderian K, Gura MA, et al. Identification of autophagy as a functional target suitable for the pharmacological treatment of mitochondrial membrane protein-associated neurodegeneration (MPAN) in vitro. *Pharmaceutics*. 2023;15:267.
63. Coe BP, Stessman HAF, Sulovari A, et al. Neurodevelopmental disease genes implicated by de novo mutation and copy number variation morbidity. *Nat Genet*. 2019;51:106-116.
64. Teunissen MWA, Lewerissa E, van Hugte EJH, et al. ANK2 loss-of-function variants are associated with epilepsy, and lead to impaired axon initial segment plasticity and hyperactive network activity in hiPSC-derived neuronal networks. *Hum Mol Genet*. 2023;32:2373-2385.
65. Sobreira N, Schiettecatte F, Valle D, Hamosh A. GeneMatcher: A matching tool for connecting investigators with an interest in the same gene. *Hum Mutat*. 2015;36:928-930.
66. Kochinke K, Zweier C, Nijhof B, et al. Systematic phenomics analysis deconvolutes genes mutated in intellectual disability into biologically coherent modules. *Am J Hum Genet*. 2016;98:149-164.
67. Indelicato E, Boesch S, Zech M. Reply to: "early onset nonprogressive generalized dystonia is caused by Biallelic SHQ1 variants". *Mov Disord*. 2023;38:1119-1120.
68. van der Spek J, den Hoed J, Snijders Blok L, et al. Inherited variants in CHD3 show variable expressivity in Snijders Blok-Campeau syndrome. *Genet Med*. 2022;24:1283-1296.
69. Harrer P, Leppmeier V, Berger A, et al. A de novo BCL11B variant case manifesting with dystonic movement disorder regarding the article "BCL11B-related disorder in two Canadian children: Expanding the clinical phenotype (Prasad et al. 2020)". *Eur J Med Genet*. 2022;65:104635.
70. Garone G, Capuano A, Amodio D, et al. BCL11B-related dystonia: Further evidence of an emerging cause of childhood-onset generalized dystonia. *Mov Disord Clin Pract*. 2024;11:897-901.
71. Yang Y, Muzny DM, Reid JG, et al. Clinical whole-exome sequencing for the diagnosis of Mendelian disorders. *N Engl J Med*. 2013;369:1502-1511.
72. Guo F, Liu R, Pan Y, et al. Evidence from 2100 index cases supports genome sequencing as a first-tier genetic test. *Genet Med*. 2024;26:100995.
73. Sanghvi RV, Buhay CJ, Powell BC, et al. Characterizing reduced coverage regions through comparison of exome and genome sequencing data across 10 centers. *Genet Med*. 2018;20:855-866.
74. Quinodoz M, Kaminska K, Cancellieri F, et al. Detection of elusive DNA copy-number variations in hereditary disease and cancer through the use of noncoding and off-target sequencing reads. *Am J Hum Genet*. 2024;111:701-713.
75. Belkadi A, Bolze A, Itan Y, et al. Whole-genome sequencing is more powerful than whole-exome sequencing for detecting exome variants. *Proc Natl Acad Sci USA*. 2015;112:5473-5478.

76. Turro E, Astle WJ, Megy K, et al. Whole-genome sequencing of patients with rare diseases in a national health system. *Nature*. 2020;583:96-102.
77. Yuan B, Wang L, Liu P, et al. CNVs cause autosomal recessive genetic diseases with or without involvement of SNV/indels. *Genet Med*. 2020;22:1633-1641.
78. Pfundt R, Del Rosario M, Vissers L, et al. Detection of clinically relevant copy-number variants by exome sequencing in a large cohort of genetic disorders. *Genet Med*. 2017;19:667-675.
79. Balasubramanian M, Willoughby J, Fry AE, et al. Delineating the phenotypic spectrum of Bainbridge-Ropers syndrome: 12 new patients with de novo, heterozygous, loss-of-function mutations in ASXL3 and review of published literature. *J Med Genet*. 2017;54:537-543.
80. Svantnerova J, Minar M, Radova S, Kolnikova M, Vlkovic P, Zech M. ASXL3 de novo variant-related neurodevelopmental disorder presenting as dystonic cerebral palsy. *Neuropediatrics*. 2022;53:361-365.
81. Jouan L, Rocheford D, Szuto A, et al. An 18 alanine repeat in a severe form of oculopharyngeal muscular dystrophy. *Can J Neurol Sci*. 2014;41:508-511.
82. Marcogliese PC, Shashi V, Spillmann RC, et al. IRF2BPL is associated with neurological phenotypes. *Am J Hum Genet*. 2018;103:245-260.
83. Frints SG, Lenzner S, Bauters M, et al. MCT8 mutation analysis and identification of the first female with Allan-Herndon-Dudley syndrome due to loss of MCT8 expression. *Eur J Hum Genet*. 2008;16:1029-1037.
84. Olivati C, Favilla BP, Freitas EL, et al. Allan-Herndon-Dudley syndrome in a female patient and related mechanisms. *Mol Genet Metab Rep*. 2022;31:100879.
85. Masnada S, Sarret C, Antonello CE, et al. Movement disorders in MCT8 deficiency/Allan-Herndon-Dudley syndrome. *Mol Genet Metab*. 2022;135:109-113.
86. Chong JX, Berger SI, Baxter S, et al. Considerations for reporting variants in novel candidate genes identified during clinical genomic testing. *Genet Med*. 2024;26:101199.
87. The UniProt C. UniProt: The universal protein knowledgebase. *Nucleic Acids Res*. 2017;45(D1):D158-D169.
88. Wiel L, Baakman C, Gilissen D, Veltman JA, Vriend G, Gilissen C. MetaDome: Pathogenicity analysis of genetic variants through aggregation of homologous human protein domains. *Hum Mutat*. 2019;40:1030-1038.
89. Shen S, Zhou H, Xiao Z, et al. PRMT1 in human neoplasm: Cancer biology and potential therapeutic target. *Cell Commun Signal*. 2024;22:102.
90. Fedoriw A, Rajapurkar SR, O'Brien S, et al. Anti-tumor activity of the type I PRMT inhibitor, GSK3368715, synergizes with PRMT5 inhibition through MTAP loss. *Cancer Cell*. 2019;36:100-114.e25.
91. Meyer E, Carss KJ, Rankin J, et al. Mutations in the histone methyltransferase gene KMT2B cause complex early-onset dystonia. *Nat Genet*. 2017;49:223-237.
92. Hashimoto M, Kumabe A, Kim JD, et al. Loss of PRMT1 in the central nervous system (CNS) induces reactive astrocytes and microglia during postnatal brain development. *J Neurochem*. 2021;156:834-847.
93. Centen LM, Pinter D, van Egmond ME, et al. Dystonia management across Europe within ERN-RND: Current state and future challenges. *J Neurol*. 2023;270:797-809.
94. Dzinovic I, Winkelmann J, Zech M. Genetic intersection between dystonia and neurodevelopmental disorders: Insights from genomic sequencing. *Parkinsonism Relat Disord*. 2022;102:131-140.
95. Peall KJ, Owen MJ, Hall J. Rare genetic brain disorders with overlapping neurological and psychiatric phenotypes. *Nat Rev Neurol*. 2024;20:7-21.
96. Indelicato E, Schlieben LD, Stenton SL, et al. Dystonia and mitochondrial disease: The movement disorder connection revisited in 900 genetically diagnosed patients. *J Neurol*. 2024;271:4685-4692.
97. Shambetova C, Klein C. Genetic testing for non-Parkinsonian movement disorders: Navigating the diagnostic maze. *Parkinsonism Relat Disord*. 2024;121:106033.
98. Baide-Mairena H, Coget A, Leboucq N, et al. Infantile-onset Parkinsonism, dyskinesia, and developmental delay: Do not forget polyglutamine defects!. *Ann Clin Transl Neurol*. 2023;10:1937-1943.
99. Herzog R, Weissbach A, Baumer T, Munchau A. Complex dystonias: An update on diagnosis and care. *J Neural Transm (Vienna)*. 2021;128:431-445.
100. Aartsma-Rus A, van Roon-Mom W, Lauffer M, et al. Development of tailored splice-switching oligonucleotides for progressive brain disorders in Europe: Development, regulation, and implementation considerations. *RNA*. 2023;29:446-454.
101. Kim J, Woo S, de Gusmao CM, et al. A framework for individualized splice-switching oligonucleotide therapy. *Nature*. 2023;619:828-836.
102. Hindorff LA, Bonham VL, Brody LC, et al. Prioritizing diversity in human genomics research. *Nat Rev Genet*. 2018;19:175-185.
103. Li S, Zhao S, Sinson JC, et al. The clinical utility and diagnostic implementation of human subject cell transdifferentiation followed by RNA sequencing. *Am J Hum Genet*. 2024;111:841-862.
104. Terkelsen T, Mikkelsen NS, Bak EN, et al. CRISPR activation to characterize splice-altering variants in easily accessible cells. *Am J Hum Genet*. 2024;111:309-322.
105. Zhao B, Nguyen MA, Woo S, Kim J, Yu TW, Lee EA. Contribution and therapeutic implications of retroelement insertions in ataxia telangiectasia. *Am J Hum Genet*. 2023;110:1976-1982.