nature reviews neurology

https://doi.org/10.1038/s41582-023-00909-9

Perspective

Check for updates

# Next-generation sequencing and bioinformatics in rare movement disorders

Michael Zech[1,2,3] & Juliane Winkelmann[1,2,4,5]

## Abstract

The ability to sequence entire exomes and genomes has revolutionized molecular testing in rare movement disorders, and genomic sequencing is becoming an integral part of routine diagnostic workflows for these heterogeneous conditions. However, interpretation of the extensive genomic variant information that is being generated presents substantial challenges. In this Perspective, we outline multidimensional strategies for genetic diagnosis in patients with rare movement disorders. We examine bioinformatics tools and computational metrics that have been developed to facilitate accurate prioritization of disease-causing variants. Additionally, we highlight community-driven data-sharing and case-matchmaking platforms, which are designed to foster the discovery of new genotype–phenotype relationships. Finally, we consider how multiomic data integration might optimize diagnostic success by combining genomic, epigenetic, transcriptomic and/or proteomic profiling to enable a more holistic evaluation of variant effects. Together, the approaches that we discuss offer pathways to the improved understanding of the genetic basis of rare movement disorders.

## Sections



[1]Institute of Human Genetics, School of Medicine, Technical University of Munich, Munich, Germany. [2]Institute of Neurogenomics, Helmholtz Zentrum München, Munich, Germany. [3]Institute for Advanced Study, Technical University of Munich, Garching, Germany. [4]Munich Cluster for Systems Neurology, SyNergy, Munich, Germany. [5]DZPG, Deutsches Zentrum für Psychische Gesundheit, Munich, Germany. e-mail: juliane.winkelmann@tum.de

## Introduction

Next-generation sequencing (NGS) has revolutionized human genetic analysis by allowing simultaneous screening for variants in hundreds of disease-related genes[1]. NGS-based massive parallelization of sequencing reactions can determine the entire nucleotide sequence of the genome of an individual in a single-analysis instrument run for less than US$ 1,000 (ref. 2). The technique is particularly suitable for molecular studies of heterogeneous Mendelian conditions, enabling specific diagnoses to be made across a wide range of phenotypes, including movement disorders[1–3].

Movement disorders encompass a vast category of neurological diseases that are frequently characterized by progressive disability[4]. Many of these conditions have an underlying genetic basis, ranging from monogenic causation to complex multifactorial aetiologies. Many subgroups exist, defined by variable expression of ataxia, chorea, dystonia, myoclonus, parkinsonism and tremor as well as phenotypically mixed syndromes, including non-movement-related symptoms, which are often individually rare and difficult to categorize on clinical grounds[4]. NGS has been instrumental in identifying the monogenic causes of rare movement disorders on a broad scale[3,5–7] and in establishing an expanding catalogue of genotype–phenotype relationships[8,9]. For example, dystonia can be related to monoallelic or biallelic variants in over 500 different genes as currently documented in the Online Mendelian Inheritance in Man database[10]. Clinicians might be unfamiliar with many of the associated diseases as they have only been reported in a few cases worldwide. In such cases, the precise molecular diagnosis can unlock important information from the literature that could be fundamental to optimizing management or offering access to disorder-specific support organizations[1,2,5].

Although many genomic data sets have been produced across different movement disorder indications since NGS became commercially available in 2011, the diagnostic rates are still capped at around 20–50%[11–15]. A major hindrance to a relevant increase in molecular aetiological yield is our inability to interpret a substantial proportion of the sequencing information that has been generated[1,2]. Individual genomes contain many thousands of ultra-rare and so-called 'private' variants, that is, sequence changes that are found exclusively in a single studied individual. Assigning clinical relevance to such variants remains challenging, even when they are discovered in known disorder-associated genes[1,2]. Limitations in the discovery power of NGS approaches are especially evident in the field of movement disorders because of marked contributions of variable expressivity and reduced penetrance as well as the high levels of allelic heterogeneity[8,9]: movement disorder-causing variants encompass a diverse spectrum of mutation types, including substitutions, deletions or duplications of single nucleotides, multi-nucleotide insertions and deletions, structural variants, repeat expansions, and mitochondrial DNA alterations[10,16,17].

To address the interpretative challenge presented by NGS, powerful computational methods have been developed to support the identification and prioritization of disease-associated variants and genes[18–21]. However, it is increasingly difficult for movement disorder specialists to oversee meaningful application of these analytics algorithms as they are being introduced at a fast pace and are not always readily usable in existing pipelines. Moreover, we recognize that the NGS diagnostic process is likely to benefit considerably from integration of other omic data, such as epigenetic signatures, transcriptomics and proteomics[22,23], into the routine care of patients with movement disorders.

In this Perspective, we highlight genomic analysis strategies and bioinformatic variant prioritization approaches in the context of rare hereditary movement disorders. We describe the principles of whole-exome sequencing (WES) and whole-genome sequencing (WGS) with a focus on the variant detection tools that are most relevant to movement disorders. We go on to discuss computational metrics and software designed to facilitate the clinically oriented filtering of variants. In addition, we outline the necessity of large-scale data sharing, including online case-matchmaking initiatives, and data integration between clinical care and research to improve diagnosis. Finally, we examine the promise of multiomics.

## Genomic sequencing and bioinformatics

WES and WGS are the main unbiased NGS techniques[1,2,24] (Fig. 1). WES targets the entire protein-coding regions (around 20,000 genes), comprising 1–2% of the human genome. This approach is highly efficient at detecting disease-associated mutations in exonic and nearby splice site sequences, which are currently thought to harbour the majority (around 85%) of known pathogenic genomic variations[2].

The diagnostic yield of WES in movement disorders has been extensively investigated across diverse phenotypes and cohorts[11]. In the broad group of ataxias, for example, 23–52% of patients were estimated to receive a specific diagnosis through exome-wide variant profiling[15,25]. WES-driven discovery has also resulted in the identification of numerous new disease-associated genes, as exemplified by the description of more than ten previously undefined monogenic aetiologies for isolated dystonia between 2015 and 2023 (refs. 26,27). Typically, the detection rates of causative variants differ according to patient characteristics in each movement disorder category, with generally higher chances of finding diagnoses in children with multisymptomatic manifestations than in adults with less complex, often multifactorial conditions[7,14,28]. However, WES has also been an invaluable tool for deciphering broad phenotypic spectra with the same genetic basis, for example, in *GNAO1*-linked conditions, in which presentations of both infantile dyskinetic encephalopathy and late-onset focal abnormal movements have been revealed[29,30].

Despite its widespread implementation as a diagnostic tool in movement disorders, WES has two key limitations[1,2]: first, inconstant depth of sequencing coverage focused on exons hinders comprehensive detection of some clinically relevant mutation types such as certain structural variants, and second, sequence alterations in non-coding DNA regions cannot be examined. These drawbacks can be overcome by adopting a WGS approach. WGS substantially reduces the likelihood of missing disease-related variants by offering uniformity of target coverage and providing analytical access to nearly all of the approximately three billion nucleotides in the genome of an individual[31]. The benefits of WGS over WES have not yet been systematically explored in the context of rare movement disorders, but pilot studies on large heterogeneous disease populations have documented diagnostic uplifts and demonstrated the effectiveness of this approach[32,33].

WES and WGS both generate extensive amounts of data (in the petabyte range for larger collections) and, therefore, require computationally sophisticated processing workflows[1,2,20]. Up to 30,000 variants can be found in an individual exome, and WGS usually yields around 3–4 million variant positions that differ from a reference genome[1,2]. Dedicated bioinformatics pipelines have an essential role in the genetic laboratory, starting with WES or WGS raw data mapping and calling of variants. The Genome Analysis Toolkit is currently the gold standard for identification of single nucleotide variants (SNVs) and short

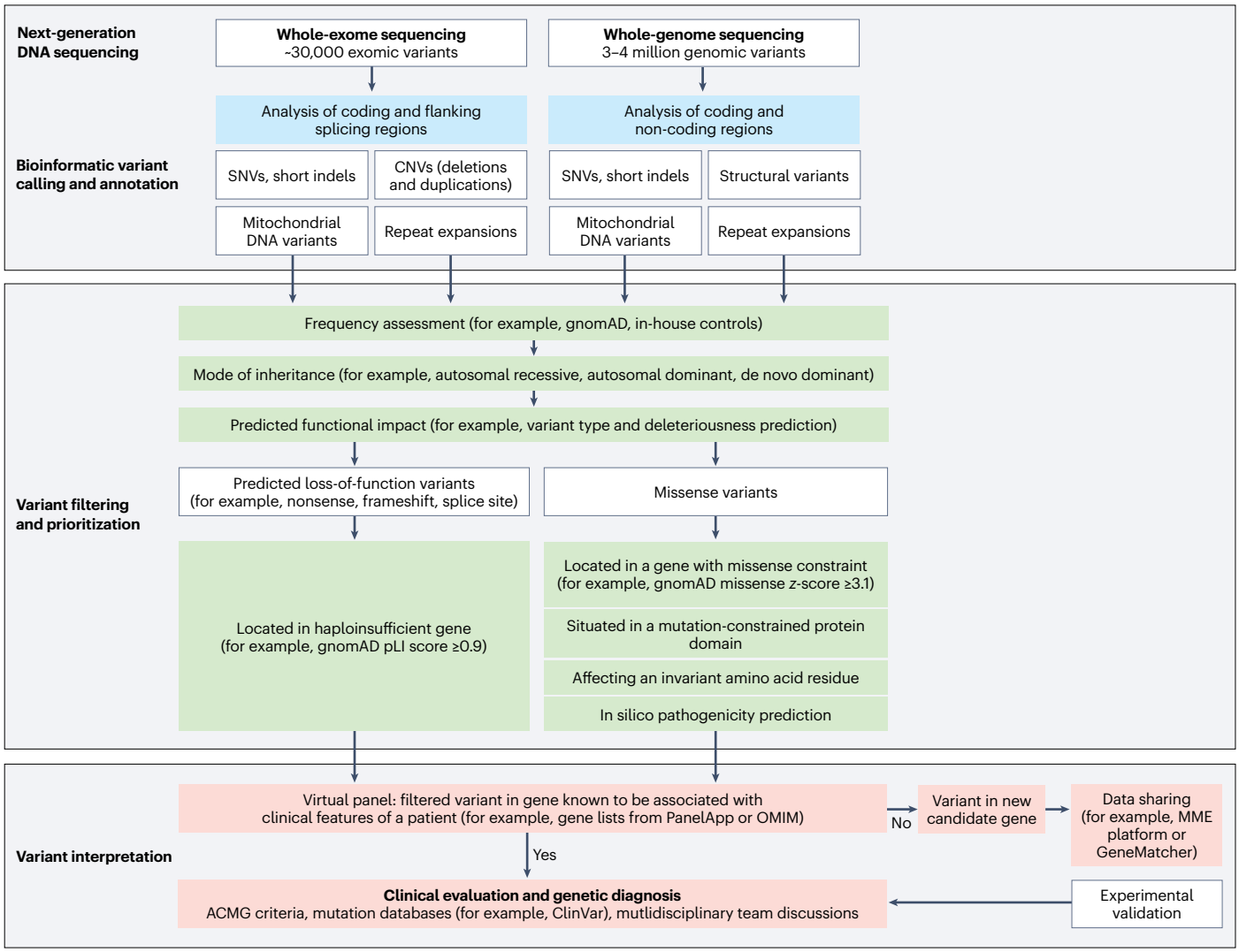


**Fig. 1 | Next-generation sequencing data production and analysis workflow.** Different variant types called from individual exome or genome raw data files are subjected to stepwise filtration, involving the integration of diverse web-based bioinformatic repositories and functional annotation sources. The filtering and prioritization steps highlighted in green are especially suitable for the analysis of single nucleotide variants (SNVs) and short insertions and deletions (indels) but might also aid the assessment of other genomic mutations such as structural variants, mitochondrial DNA variants and repeat expansions. The workflow has some technical limitations that can hinder reliable detection of certain genotypic abnormalities, such as single-exon copy number variations (CNVs), as well as larger or non-coding repeat expansions. Once a shortlist of candidate genetic alterations has been identified, an expert review is conducted to allow the determination of rare variants that have high levels of evidence for association with the disease. ACMG, American College of Medical Genetics and Genomics; gnomAD, Genome Aggregation Database; MME, Matchmaker Exchange; OMIM, Online Mendelian Inheritance in Man; pLI, probability of being loss-of-function intolerant.

(1–50 bp) insertions and deletions (indels)[20,34], which represent the majority of recognized movement disorder-causing mutations. Several NGS studies of patients with ataxia, dystonia, mixed hyperkinetic syndromes or other rare movement disorders have shown that SNVs and indels account for the majority (85–95%) of molecular diagnoses[6,12,14,15], although these numbers might be biased owing to incomplete assessment of other variant types, especially in earlier work.

WES and WGS data can also be exploited to assess copy number variations (CNVs), for which an increasing battery of detection algorithms is being developed[35,36]. Tools that identify CNVs in short reads from NGS machines can be coverage-based callers (for example, ExomeDepth or CNVnator) or callers using integrated paired-end and split-read analysis strategies (for example, DELLY or Manta)[37,38], which can detect deletion and duplication events with high sensitivity and specificity. CNVs represent a relevant class of genomic alterations that contribute to movement disorder manifestations, and with the growing adoption of CNV screening tools in NGS pipelines, we are identifying intriguing 'new' roles for 'old' deletion syndromes in movement disorders such as 22q11.2 microdeletions (DiGeorge syndrome) in parkinsonism and hyperkinetic phenotypes[39,40]. Importantly,

## Glossary

**Coverage-based callers**
Copy number variant detection tools that determine the presence of a deletion or duplication by comparing the read coverage in the affected genomic interval with the rest of the sequenced exome or genome. Higher sequencing depth is necessary for reliable analysis.

**Digenic inheritance**
A mechanism whereby the expression of a disease phenotype is determined by the presence of genetic pathologies in two different loci, often associated with epistatic interactions between these loci (encoded proteins might act in the same pathway).

**Generative artificial intelligence**
Algorithms that can be used to produce new content, including synthetic data.

**Integrated paired-end and split-read analysis strategies**
Paired-end mapping approaches can define copy number variants on the basis of alterations in the insert size of paired-end reads, whereas split-read approaches are helpful for predicting copy number changes by assessing unaligned discordant reads that were split and mapped separately from the reference genome.

**Mapping certainty**
A measure of the accuracy of alignment of sequencing reads to the correct location in the genome. Can be confounded by DNA characteristics such as repetitive regions.

**Massive parallelization**
A high-throughput approach used in next-generation sequencing studies, which allows analysis of millions of short reads (usually containing 100–150 bp) in an automated miniaturized fashion. This approach differs from traditional capillary Sanger analysis in terms of time-effective mass production of sequencing outputs.

**Mendelian conditions**
Clinical diseases that are caused by high-effect rare variants in single genes, in contrast to polygenic or multifactorial diseases, which are associated with many common variants with low effect sizes at various genomic loci and are influenced by other non-genetic factors.

**Missense constraint**
A measure of genetic intolerance to amino acid substitutions, which can aid prioritization of gene candidates involved in missense mutation-associated diseases.

**Mobile element**
Genomic sequences that can move between chromosomes, for example, through cut-and-paste mechanisms in DNA transposons. These elements have a role in genome evolution, and their integration into disease-associated genes can disrupt the open reading frame and cause clinical phenotypes.

**Penetrance**
A measure of the proportion of carriers of a specific monogenic disease predisposition who present with clinical features of the associated condition.

**Phenotypic pleiotropy**
A phenomenon whereby variants in a disease-related gene are associated with multiple (similar or divergent) phenotypic abnormalities.

**Simplex cases**
Individuals with a disease phenotype who have no relatives affected by the same condition.

**Spike-in panel**
A protocol that dynamically incorporates specific DNA segments into the sequencing analysis; for example, complementary interrogation of all base pairs of the mitochondrial genome in addition to the nuclear coding sequences in the form of a mitochondrial spike-in panel in diagnostic exome studies.

WGS offers the opportunity to screen genomes for a range of structural variations beyond those identifiable by WES[41]; for example, WGS has been successful in uncovering a pathogenic inversion disrupting *QDPR*[42] (encoding dihydropteridine reductase), a gene that has been implicated in tetrahydrobiopterin deficiency-related movement disorders. Most CNV-calling algorithms have limited efficacy for the discovery of small CNVs, particularly those containing only one or two exons. These events are identified as single aberrant signals in the data with often suboptimal intensity and unwanted noise, creating challenges in routine clinical applications owing to false negatives and false positives[43].

Another important development for WES and WGS data analysis in patients with movement disorders is the introduction of methods that allow scrutiny of mitochondrial DNA mutations and pathological repeat expansions[44–46]. With respect to mitochondrial DNA mutations, so-called 'off-target reads' generated during standard WES or WGS experiments and processed in the bioinformatics pipeline can be reliably utilized for molecular diagnosis[44]. Bespoke processing workflows in genetic laboratories can integrate off-capture sequencing results that derive from the enrichment of DNA fragments outside the intended target regions, including the mitochondrial genome[47]. A retrospective evaluation of 11,424 WES data sets reported the detection of pathogenic mitochondrial DNA variants in 11 individuals, including patients with ataxia, dystonia or myoclonus[48]. Despite their practical applicability, off-target read-based mitochondrial DNA variant screens can only be performed in WES or WGS studies to determine diagnoses when the capture-target kit supports the interrogation of the mitochondrial genome (for example, in the form of a spike-in panel with the core nuclear exome)[49].

With respect to pathological repeat expansions, a specific tool named ExpansionHunter Denovo[50] is gaining popularity. This tool is suitable for performing hypothesis-free, genome-wide repeat profiling and has high diagnostic accuracy. In families with late-onset cerebellar ataxia, a comparatively prevalent but often genetically intractable syndrome, application of this tool pinpointed a new repeat expansion disorder—an adult-onset cerebellar ataxia syndrome known as ATX-*FGF14* (ref. 51). Furthermore, a systematic assessment of the test performance of ExpansionHunter-supported repeat expansion profiling for 13 neurological disorders caused by these mutation types showed that WGS could distinguish between expanded and non-expanded alleles with 97.3% sensitivity and 99.6% specificity[46]. By contrast, WES has restricted ability to find causative expanded sites in patients with repeat expansion-associated movement disorders because this method cannot accurately calculate the size of alleles larger than the commonly used read length of 100 bp and cannot capture non-coding parts of disease-related genes (for example, intron 1 of the *FXN* gene (encoding frataxin), which harbours the Friedreich ataxia-linked GAA repeat)[2]. WES-based molecular diagnostics for movement disorders might therefore require additional testing for pathogenic repeats on alternative platforms, depending on the phenotypic characteristics of the examined patient.

Other rare complex mutational events underlying monogenic movement disorders have begun to be unravelled using NGS data and

modern analytical methodologies, including mobile element insertions in *NKX2-1*-linked childhood-onset chorea[52]. In daily practice, genetic data analysts greatly benefit from the simultaneous integration of multiple independent variant callers in their pipelines, and these systems should ideally offer the possibility to incorporate newly emerging tools once their diagnostic sensitivities and specificities have been validated.

## Variant prioritization and pathogenicity assessment

Genomic sequencing, in particular WGS, produces an abundance of variant data, posing a challenge to determine which of the sequence changes have a causal role in the phenotype of an individual[1,2]. To address this diagnostic bottleneck, automated workflows have been developed, which support analysts in the assessment of the pathogenic role of variants. The process follows a series of computational mutation-filtration steps that depend heavily on a range of online resources and bioinformatics tools[18–20] (Fig. 1).

A primary filtering strategy involves cross-referencing of patient-derived variants to catalogues of variants that are found in population controls; this step helps to filter out benign alterations that are observed in individuals who are not affected by the disease[53,54]. The Genome Aggregation Database (gnomAD), which contains information from over 120,000 exomes and over 15,000 genomes of various geographical origins, is commonly used for this purpose[21]. In the context of movement disorders, an important caveat needs to be considered when deploying variant exclusion with gnomAD data (Box 1), namely, that the data set is not depleted for alleles associated with adult neurological conditions and reduced penetrance[1]. For example, the pathogenic p.Glu303del variant in *TOR1A* (encoding torsin 1A), which causes autosomal dominant generalized dystonia[27] (penetrance around 30%), is present in 30 heterozygous gnomAD carriers[21]. Moreover, we increasingly observe that variants linked to newly discovered autosomal recessive movement disorder phenotypes, such as specific *WARS2* (encoding tryptophan–tRNA ligase, mitochondrial) or *SHQ1* (encoding protein SHQ1 homologue) mutations in parkinsonism and myoclonus[55,56], are found in a homozygous state in gnomAD participants[21], highlighting the need for careful literature-informed evaluation. Additional web-based curated reference catalogues exist for CNVs, including the Database of Genomic Variants[57] and dbVar[58], which host structural variation data from healthy populations (Database of Genomic Variants) or both controls and individuals with clinical phenotypes (dbVar).

The inheritance of variants, ideally based on recognizable familial transmission patterns, should also be included in the filtering criteria[53,54]. For simplex cases, parent–patient trio analysis has proved to be a highly efficient strategy to reduce the analytical burden by enabling straightforward detection of de novo variants through a bioinformatics 'subtraction' approach[59]. Studies show that de novo mutational events, including recurrent hits observed in multiple patients (for example, *ADCY5*-related dyskinesia)[60], constitute a major cause of early-onset movement disorders[7,14]. Of note, new challenges in the filtering of variants based on documented inheritance modes in movement disorders arise from the observation that a growing number of genes are being linked to both dominant and recessive phenotypes[61].

A further step is to prioritize variants according to their functional consequence and assumed deleteriousness[53,54], using powerful computational metrics and software packages[19,20]. Currently, it is advisable to focus on protein-changing variants, which consist of two broad classes: predicted loss-of-function (LoF) variants (that is, nonsense, frameshift and splice site alterations) and missense variants. In theory, LoF variants should be excellent candidates for disease causation because they are expected to disrupt the reading frame of the gene. However, these variants are abundant in the population and each genome carries around 100 such alterations[2].

## Box 1

# Limitations of bioinformatic metrics and online tools

Several limitations must be considered when analysing variants associated with rare monogenic movement disorders.

### Genome Aggregation Database (gnomAD)

- Owing to reduced penetrance, pathogenic or likely pathogenic variants linked to autosomal dominant movement disorders might be found among unaffected gnomAD participants.
- Homozygous carriers of certain pathogenic or likely pathogenic variants linked to autosomal recessive disorders might also be found among unaffected individuals because these mutations represent hypomorphic alleles; examples include variants in *WARS2* and *SHQ1*.
- Loss-of-function variants in genes with low scores in the probability of being loss-of-function intolerant can be causative for rare movement disorders owing to reduced penetrance or imprinting phenomena; examples include variants in *VPS16* and *SGCE*.
- Missense variants in genes with low missense *z*-scores can be causative for rare movement disorders if the gene has high rates of benign missense variation; examples include variants in *ANO3*.

### Local missense constraint metrics

- Missense variants located outside conserved domains or constrained regions can be causative for rare movement disorders owing to their site-specific mutational effects.

### In silico pathogenicity predictions

- Variants evaluated as benign can be causative for rare movement disorders, for example, when they occur at amino acid positions with poor evolutionary conservation; examples include variants in *PRKRA* and *EIF2AK2*.

### GeneMatcher

- Most GeneMatcher entries refer to paediatric cases with syndromic diseases and therefore it is often more difficult to find matches for candidate genomic findings from patients with rare movement disorders.

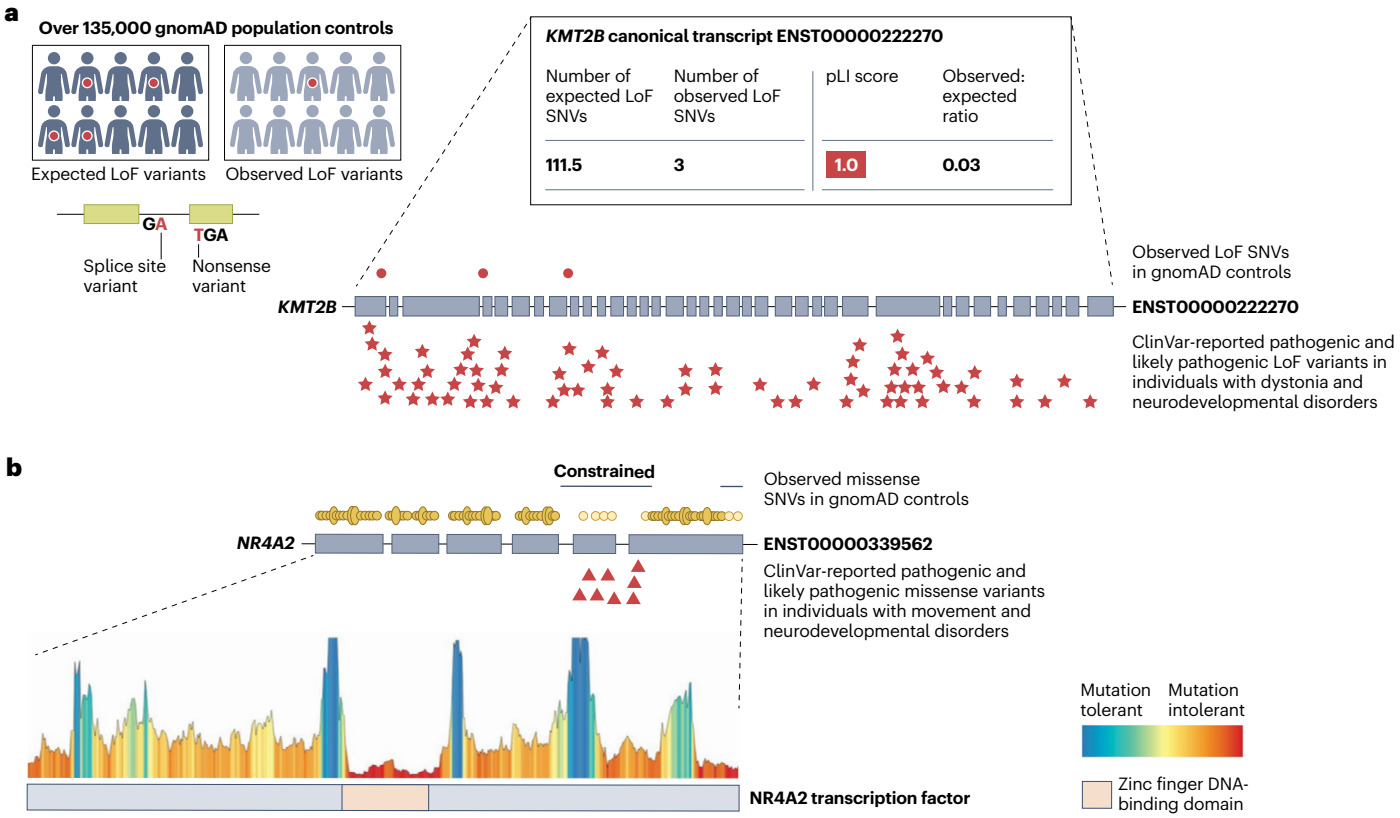


**Fig. 2 | Mutational constraint metrics to aid variant interpretation.** **a**, Sequencing information from over 120,000 exomes and over 15,000 genomes in the Genome Aggregation Database (gnomAD) is used to provide constraint scores for a given mutation type such as loss-of-function (LoF) variation, for example, nonsense and splice site mutation-inducing single nucleotide variants (SNVs)[21]. The probability of being LoF intolerant (pLI) metric (range 0.0–1.0) is calculated for each gene in gnomAD based on the number of observed versus expected rare LoF SNVs, taking into account the length and nucleotide sequence of the gene. Genes with pLI scores ≥0.9 are considered to be under severe constraint against LoF mutations. The figure shows the LoF variant-constrained gene *KMT2B* (encoding histone–lysine *N*-methyltransferase 2B), mutations of which cause dystonia 28, childhood-onset[62] (Online Mendelian Inheritance in Man #617284). In gnomAD controls, *KMT2B* has significantly fewer LoF SNVs than expected (pLI score 1.0), indicating a high degree of evolutionary selective pressure. Consistent with this finding, heterozygous LoF *KMT2B* variants are responsible for highly penetrant paediatric dystonia syndromes; the distribution of such mutations registered in ClinVar as of May 2023 is depicted below the *KMT2B* transcript[17]. **b**, The degree of regional missense mutation constraint can also be estimated using gnomAD data[21,66,67]. For example, *NR4A2* (encoding nuclear receptor subfamily 4 group A member 2), a gene linked to a neurodevelopmental disorder with dystonia and parkinsonism (Online Mendelian Inheritance in Man #619911)[69–71], contains a coding sequence with significantly fewer missense variants than expected. Missense variants that map to the area with local missense intolerance might be regarded as high-priority candidates for disease causation. In ClinVar, disease-related missense mutations cluster within this region[17], which encodes a functionally important protein domain. Computational tools, such as the MetaDome web server[67], offer user-friendly visualization of missense-constrained protein regions inferred from gnomAD data as illustrated in the bottom panel. Specific pathogenic and likely pathogenic *NR4A2* missense variants are shown based on ClinVar data accessed in May 2023 (ref. 17).

To distinguish genes in which LoF variants are tolerated from those that are LoF intolerant, the gnomAD data-based 'probability of being LoF intolerant' (pLI) score has been introduced[21] (Fig. 2a). This metric, which is calculated based on a comparison of observed versus expected LoF variants for each gene in gnomAD participants, provides a statistically robust method for the prioritization of LoF mutations that are likely to be clinically relevant[21] and has facilitated the establishment of many genotype–phenotype associations in patients with movement disorders[14]. The pLI score has aided the discovery of *KMT2B* (encoding histone–lysine *N*-methyltransferase 2B) haploinsufficiency as a cause of childhood-onset dystonia[62] and has assisted in the identification of genes in which LoF variants are generally considered to be irrelevant to movement disorder traits (for example, LoF variants in *LRRK2* (encoding leucine-rich repeat serine–threonine-protein kinase 2) do not underlie hereditary Parkinson disease[63,64]).

For missense variants, a similar measure, the gnomAD-derived missense *z*-score[21], is available. This score allows filtering of WES and WGS data for genes in which missense variants are significantly under-represented among control individuals. Missense substitutions in such genes should be carefully evaluated as the probability of pathogenicity is likely to be high. A typical example of a movement disorder-related gene with severe missense constraint is *ATP1A3* (ref. 21) (encoding sodium–potassium-transporting ATPase subunit α3), which is linked to dystonia–parkinsonism and infantile dyskinetic syndromes[65]. Constraint-based approaches for missense variant prioritization have also been developed at the levels of protein domains

and individual genomic positions (that is, at the codon level)[66–68] where such mutations can occur (Fig. 2b). These computational methods exploit the fact that certain regions of genes and their products show mutational invariability in the general population[66,67]. It is possible to specifically screen for missense variants that map to these mutation-intolerant sites and are, thus, more likely to have a deleterious effect. For example, in *NR4A2*-associated neurodevelopmental disorder with dystonia and parkinsonism[69–71], nearly all pathogenic missense variants are located in a protein motif that shows minimal functional variation in population controls[72]. Box 1 highlights limitations in the use of mutational constraint parameters for the interpretation of LoF and missense variant pathogenicity in rare movement disorders.

Additional information for filtering missense variants includes predicted effects on protein structure and evolutionary conservation, both of which can be assessed by several commonly available in silico classifiers[19,20,73]. Some recently introduced tools combine multiple outputs from different published algorithms to estimate the functional deleteriousness of a given amino acid change[74]. These so-called 'metapredictors' can assist with missense variant evaluation and generate high positive predictive values, but their results need to be interpreted in conjunction with the peculiarities of the relevant disease-associated proteins (Box 1).

An additional stage, which is useful when variants are pre-filtered as described above, involves the implementation of candidate gene lists obtained based on the phenotypic characteristics of a patient[53,54]. This 'virtual panel' approach narrows the list of selected variants to those affecting genes that have an established association with the presenting clinical features. In 2019, a regularly curated virtual gene panel catalogue known as PanelApp was launched through a publicly available platform[75]. Alternatively, gene lists can be downloaded from Online Mendelian Inheritance in Man[10], although not all entries might be up to date.

For any prioritized variant, a standardized framework for clinical interpretation, such as the five-tier classification system provided by the American College of Medical Genetics and Genomics, must be applied[76]. To reduce inter-rater variability and the risks of subjective evaluation of variants, software tools that provide automated American College of Medical Genetics and Genomics guideline-based interpretative outputs for filtered variants, such as InterVar[77], are becoming available.

Despite considerable advances in bioinformatics-driven variant categorization, many sequence changes remain of undetermined clinical significance (so-called variants of uncertain significance (VUSs)). A major controversy is whether variants for which a consensus cannot be reached on disease causality should always be reported back to referring clinicians and the affected families[78]. The classification criteria for VUSs are conservative and designed to prevent potential harm that could result from erroneous pathogenicity assignments based on insufficient evidence[76]. Reporting of these variants can trigger manifold patient responses and should follow careful guidelines to avoid enhanced medical uncertainty and negative psychosocial impact[79].

One important strategy for dealing with VUSs is to reanalyse unresolved WES and WGS data at periodic intervals[80]. Reanalysis of existing NGS data, including integration of updated database annotations, consideration of more detailed phenotype information, and searches for the latest published gene–disease and variant–disease relationships, was shown to increase the diagnostic yield by 6–47% through various measures, including VUS upgrading[81]. However, VUS reclassification through reanalysis might be difficult to achieve for patients from understudied geographical areas who display distinct allelic architectures with specific rare variants, including founder mutations that are not registered in available reference databases. Global efforts are needed to generate ancestry-specific allele data sets as conducted for a large population from the Middle East[82]. As WGS becomes more widely deployed in the field of monogenic movement disorders, we are likely to see exponential growth in the number of difficult-to-interpret variants in non-coding genomic regions, which will further increase the quantity and complexity of VUS information.

Identification of variants that are unequivocally causal for movement disorders remains a difficult challenge because of the genetic heterogeneity and clinical variability associated with these conditions. In an example of extreme phenotypic pleiotropy, independent groups have implicated genes of the nucleotide excision DNA repair pathway, which were previously linked to hereditary skin disorders with photosensitivity and cancer, in rare movement disorders characterized by chorea, dystonia and/or ataxia[83,84].

## Data sharing

Over the past few years, it has become clear that genomic data produced by individual laboratories are frequently insufficient to generate compelling evidence for causality of VUSs in known disease-associated genes or candidate variants in potential, new disorder-relevant loci[85]. In rare movement disorders, only a very few individuals seen at a single institution are affected by a specific syndrome, and the vast majority of sequenced individuals in each local database do not share the same variant hits. Establishment of mutation recurrence in independent similarly affected individuals and identification of multiple patients with variants in the same gene are necessary to firmly define genotype–phenotype correlations[86].

The challenge posed by VUSs and the situation of having identified one single family with a promising but unconfirmed gene candidate — termed the '*n*-of-1 problem' — can be overcome by sharing observations regarding rare molecular findings with the broader genetics community[85,87]. Collaborative efforts in global projects have addressed this need by developing data-sharing solutions that provide centralized repositories for clinically evaluated variants as well as platforms for genotype–phenotype matchmaking[87,88]. Databases that actively catalogue sequence changes according to their previously reported disease relevance include ClinVar[17] and the Human Gene Mutation Database[89]. These sources summarize information on variant pathogenicity, mostly for SNVs and short indels, that would otherwise be dispersed across the literature and private mutation compendia, allowing analysts to quickly judge the pathogenic significance of variants identified through WES or WGS.

Similar clinically centred annotation platforms are available for CNVs (DECIPHER)[90] and mitochondrial DNA variants (MITOMAP)[91]. However, as these databases are human-curated and sometimes filled with spurious genotype–phenotype associations, misclassifications or conflicting interpretations of variants are not uncommon[1]. This situation is especially problematic in the molecular analysis of rare movement disorders because reliably interpreted variant calls are generally under-represented in these indications (as compared with, for example, intellectual disability), which can increase the burden of diagnostic uncertainty. Therefore, genetic laboratories that focus on movement disorders should be urged to systematically submit their sequencing results to ClinVar and other public knowledge repositories. Alternative community-based curation platforms, such as the Movement Disorder Society Genetic mutation database, have

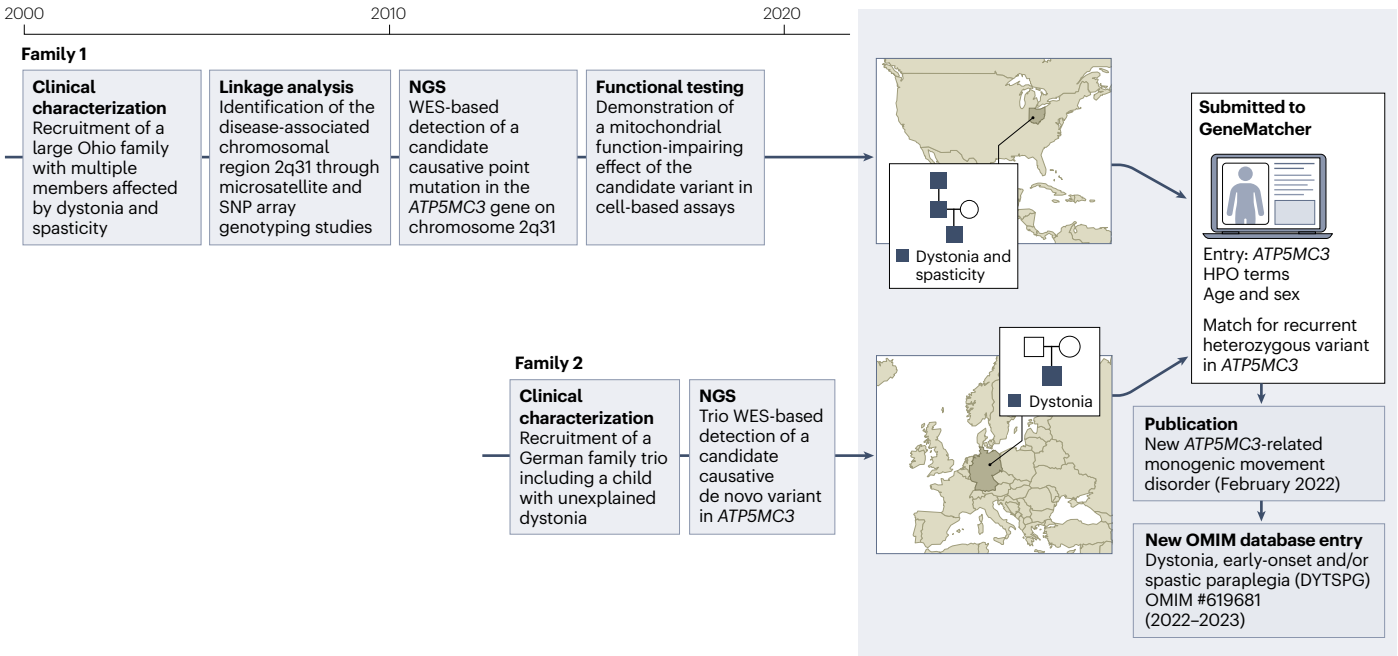


**Fig. 3 | Case matchmaking and disease gene discovery via the GeneMatcher platform.** A candidate ultra-rare variant in the mitochondrial complex V gene *ATP5MC3* was detected in a large pedigree affected by dystonia and spasticity from the USA state of Ohio[120]. Despite functional molecular characterization of the variant, the definition of a new hereditary disorder was not possible for more than a decade because no additional independent case with the same genetic defect had been identified. In 2019, the Ohio family was 'matched' through the GeneMatcher[95] node to a German patient who harboured an identical *ATP5MC3* mutation. Consequently, the *ATP5MC3*-related monogenic movement disorder became firmly established[97]. HPO, Human Phenotype Ontology; NGS, next-generation sequencing; OMIM, Online Mendelian Inheritance in Man; SNP, single nucleotide polymorphism; WES, whole-exome sequencing.

been launched to promote meaningful exploration of the evidence of variant pathogenicity in the context of ataxia, chorea, dystonia and other movement disorder presentations[92]. Nevertheless, a lack of diversity in genomic testing among ethnic groups, with marked under-representation of certain populations (for example, minority groups and people with African ancestry), remains a major hurdle for data sharing and interpretation[93].

To increase the analytical power of rare disease diagnostics, important genotype-driven and phenotype-driven matching algorithms have been established, including the international Matchmaker Exchange (MME) service[94]. Initiated in 2013, MME introduces genetic data-sharing mechanisms and tools for phenotypic analysis that are incorporated into a federated system with the goal of catalysing connections between clinicians and researchers with an interest in the same genes and disorders[94].

The MME network is joined by a series of connected nodes, among which GeneMatcher is one of the most widely applied tools[95]. GeneMatcher includes data from thousands of individuals with rare disease phenotypes, and is accessible to medical professionals from around the globe, thereby facilitating the identification of patients with similar genotypic and phenotypic profiles[95] (Fig. 3). Using this platform, a substantial number of cases of rare and ultra-rare movement disorders have been matched, leading to characterization of many previously unrecognized disease entities[96]. For example, WES recently revealed a private missense variant in *ATP5MC3* — a gene that encodes an essential component of the mitochondrial respiratory chain complex V — in a family in the USA affected by dominantly inherited dystonia and spasticity[97]. Because the variant had never been described before in independent cases, it qualified as a VUS and the family remained undiagnosed. The finding was entered into GeneMatcher, which ultimately yielded a match through identification of the same mutation in a German dystonia pedigree, resulting in the discovery of a novel mitochondrial defect-related monogenic movement disorder[97].

Several additional data-sharing initiatives that allow comparison of sequencing findings, such as platforms that register systematic information on de novo variants identified from trio WES and WGS analyses[98], can aid genetic studies of rare movement disorders. The de novo variants are mostly derived from patients with neurodevelopmental diseases[98], but their consideration can be useful in movement disorder diagnostics given that movement disorders and neurodevelopmental diseases often share a common genetic basis[14]. Interoperable national and continent-wide data hubs, which offer improved pathways to the sharing of ethnically diverse genetic information in common databases, are also under development[85]. For example, the European Genome–Phenome Archive and its German hub, the German Human Genome–Phenome Archive, support the deposition of genomic sequences and phenotypes, including movement disorders, to optimize data reuse and accelerate disease-associated gene discovery[99]. Similarly, efforts are being geared towards the sharing of pan-European rare disease data in a systematic manner within the Solve-RD research consortium[100]. Importantly, when data sharing through health professionals is unable to provide diagnostic clarity, some patient families take responsibility and advertise their genetic information on social media to make themselves more 'discoverable'[87].

## Integration of multiomic studies

DNA-level approaches, such as WES and WGS, are limited in their capacity to clarify the significance of a large proportion of variants in disease manifestation[1,2]. Although re-analyses, predictive algorithms and data sharing have improved the prioritization and interpretation of genetic findings, these methodologies are often unable to confirm or refute the pathogenicity of VUSs and alterations situated in non-coding regions. Parallel assessment of additional layers of omics offers an opportunity to overcome these hurdles[22] (Fig. 4). A growing body of literature is linking genomic sequencing results with epigenetic, transcriptomic and/or proteomic data to reveal pathophysiological mechanisms and uncover diagnoses in previously unresolved monogenic phenotypes[23].

At present, strategies for multiomic data integration are not being systematically applied to rare movement disorders but early studies demonstrate their promise for improving diagnostic performance[101], and international collaborations, such as the European Joint Programme on Rare Diseases, have been put in place to scale their use in molecular characterization of patients with dystonia and other indications. Genome-wide analysis of DNA methylation marks can identify biologically meaningful signals that might support the evaluation of variant effects[102]. Initially introduced in the field of neurodevelopmental diseases, these genomic episignatures were shown to be especially useful for investigating conditions linked to genes that are suspected to influence DNA methylation status[102].

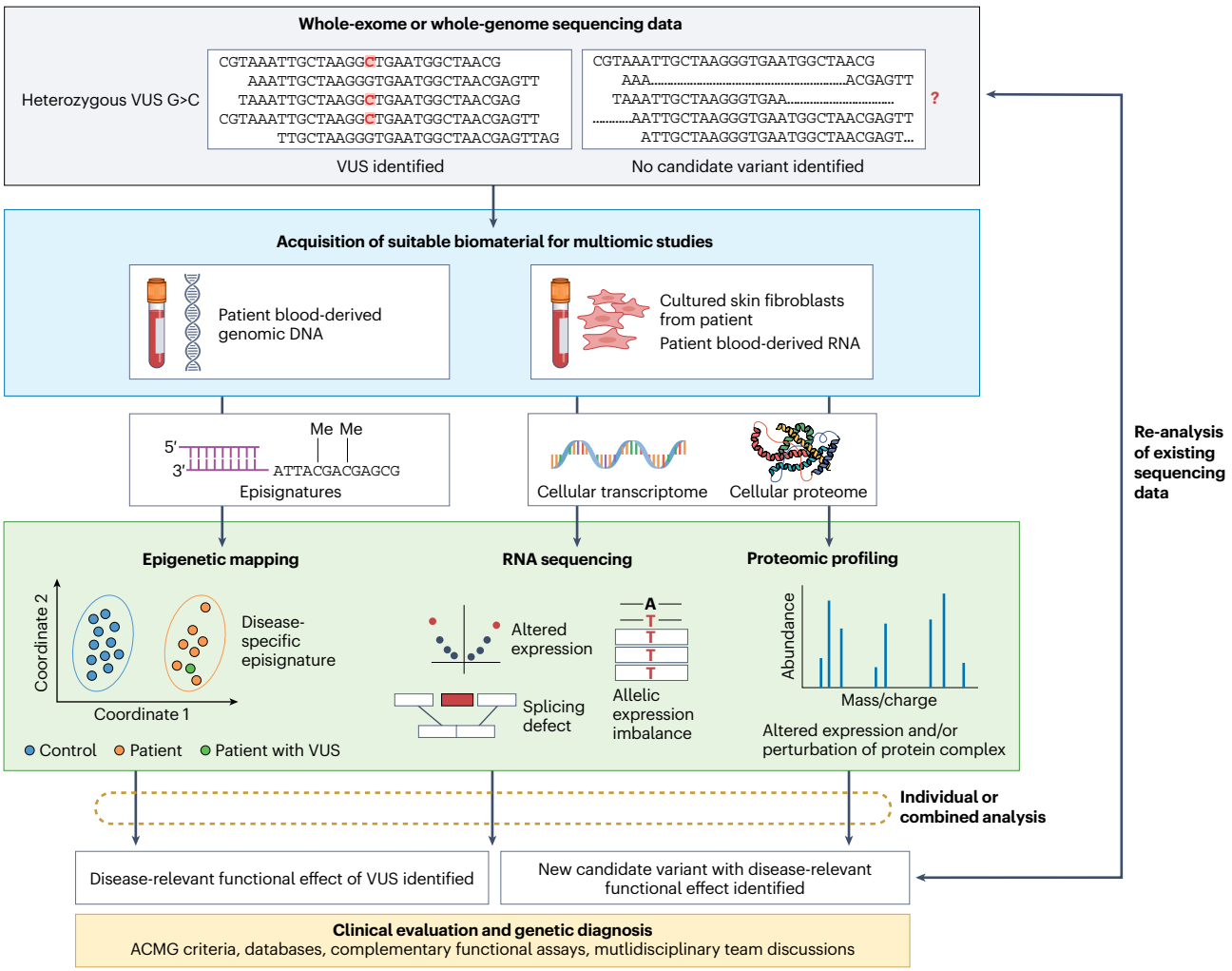


**Fig. 4 | A suggested multiomic-based diagnostic strategy.** Multiple components of disease-causing molecular lesions, including DNA-level aberrations, disorder-specific DNA modification patterns, and alterations in RNA and protein expression, can be considered in integrated multiomic analyses. DNA methylation marks, also known as episignatures, are emerging as powerful tools to characterize the significance of variants in relation to rare phenotypes, especially those linked to defects of the epigenetic machinery. RNA sequencing (RNA-seq) can detect different types of variant-induced transcript pathologies, including aberrant expression, defective splicing and monoallelic expression states. Quantitative proteomics can be used to find protein expression outliers caused by aetiologically involved variants and to characterize associated protein complex disturbances. 'Standard' peripheral blood-derived DNA is suitable for the study of episignatures but other patient-specific biological samples might be more useful for other analyses such as RNA-seq or proteomics. Skin fibroblast cultures have been found to represent an optimal biomaterial for such multiomic approaches[107]. Alternatively, whole-blood RNA-seq can be a robust strategy for the profiling of disease-relevant transcript expression and splicing defects in patients with monogenic diseases[108]. The different analytical dimensions of multiomic tests can be assessed separately or in parallel to maximize molecular diagnostic yields. ACMG, American College of Medical Genetics and Genomics; VUS, variant of uncertain significance.

Studies have been undertaken to develop accurate episignature-based classifiers for variants in the dystonia-linked gene *KMT2B*, which encodes a histone methyltransferase involved in epigenetic modifications[103,104]. In one publication, a blood-derived, disorder-specific episignature on 113 DNA methylation sites was used to re-classify four VUSs in *KMT2B*, three of which newly qualified as disease causing, leading to optimization of diagnostic outcome and therapy-relevant results, given that *KMT2B*-related dystonia is highly responsive to deep brain stimulation[103]. Moreover, DNA methylation profiling seems to enable predictions of age of onset and disease severity in patients with *KMT2B* mutations[103].

The study of transcriptomes by RNA sequencing (RNA-seq) represents another complementary assay to WES and WGS analyses[105]. This technique examines RNA levels in an unbiased manner, both qualitatively (integrity of transcripts) and quantitatively (expression levels), and can provide a broad view of transcription-related pathological events[105]. RNA-seq data are particularly important for the interpretation of non-coding variations but can also be used to assess the effects of synonymous variants that can affect splicing[22]. Most RNA-seq pilot studies performed on patients with heterogeneous rare disease presentations have aimed to decipher the roles of uncertain WES or WGS findings by focusing on the detection of splicing mutation-induced aberrant transcripts and/or aberrant expression states[105,106]. The potential diagnostic value of this approach in individuals with movement disorder features was demonstrated by a large-scale combined WES–WGS–RNA-seq study[107], which described missplicing and pathologically decreased expression of the complex I assembly factor gene *TIMMDC1* as a result of deep intronic variants in patients with ataxia and dyskinetic movements.

In multiomic studies, further diagnostically useful information can be derived from proteomic investigations[22]. Proteomics assist in rare disease variant interpretation by identifying instances where VUSs have resulted in abnormal upregulation or downregulation of protein expression. Though still in its infancy in rare movement disorder diagnostics, a quantitative proteomic approach was recently applied successfully in a child with dyskinetic epileptic encephalopathy for functional validation of candidate variants in *ATP5PO* (encoding ATP synthase subunit O, mitochondrial), not only establishing the diagnosis but also characterizing a new recessive neurogenetic syndrome[101].

Tissue-specific expression is an important aspect that needs to be considered in both transcriptomic and proteomic studies[23]. Brain tissue is rarely available for such studies, and samples are routinely extracted from patient-derived skin fibroblasts, in which thousands of RNAs and proteins can be reliably assessed[107]. An alternative approach is to investigate blood transcriptomes, which allows non-invasive diagnostic identification of RNA defect-related molecular drivers of disease[108]. Ideally, the results of the different omic analyses should not be evaluated separately but should be processed through a unifying bioinformatic framework, or multiomics pipeline, which allows superimposition of all layers of information to maximize power for functional annotation of variants. Further omic methods beyond those described above — for example, metabolomics — might also be incorporated into diagnostic workflows for rare movement disorders.

## Conclusions and future opportunities

The advent of NGS with its associated computational analytical tools has opened up a new era in diagnostics for rare movement disorders. WES and WGS will undoubtedly become the cornerstone for molecular analysis of most if not all patients affected by these conditions. In this scenario of broad application across diverse disease subgroups, it will be vital to establish broadly accepted standards for incorporation into daily routine and interdependent training of movement disorder specialists and neurogenetics experts, who should work together to implement clinically meaningful genomic medicine to benefit patients with rare movement disorders (Box 2).

Evidence-based diagnostic pathways towards a genomic analysis-first approach must be developed for individual movement disorder indications. For example, in the field of dystonia, a predictive clinical scoring system has been proposed that incorporates genomics into routine care[14]. Moreover, improvements are needed in the analysis of technically difficult-to-identify mutations, VUS assessment and the integration of information from multiple omics sources to realize the full potential of large genomic data sets[87,88]. Concerted efforts

## Box 2

### Incorporating genomic-driven diagnostics into clinical practice

Genome-informed care is emerging as a major focus for patient-centred precision medicine in the field of rare movement disorders and, therefore, a roadmap for inclusion of transformative technological advances in clinical practice is needed. The question of how best to design coherent clinical implementation models across different national and international health systems remains unresolved. Human geneticists provide expertise in the analysis and primary interpretation of genomic data, whereas movement disorder neurologists serve as specialists for the clinical assessment of affected families.

In an optimal scenario, physicians from the various movement disorder specialty areas would be directly trained in the everyday implementation of genomic-guided patient evaluation and interpretation of genetic results. Standardized knowledge-building programmes would be desirable for clinicians who choose to specialize in this domain, providing guidance on how to consider diagnostic yields and technological limitations, assess the pathogenicity status of genomic findings, and communicate the meaning of relevant identified variants to families. However, current health-care policies in many countries favour general neurological care over specialty care, even in major academic centres, thereby hampering the development of practice strategies in which physicians can act as both movement disorder and neurogenetic experts. In these settings, close collaboration and interdisciplinary case conferences will be essential, engaging teams of interacting movement disorder neurologists, laboratory genomics specialists and clinical genetics analysts.

Clinicians in the rare movement disorders community cannot afford to ignore the unprecedented insights that genomics can provide into the aetiologies of the conditions that they treat. We have aimed to highlight the necessity of understanding the processes and diagnostic importance of new genomic technologies, envisaging a future in which health-care providers can integrate these technologies into clinical workflows and support the education of dedicated genomics specialists.

of the neurogenomics community, which is composed of clinicians, human geneticists, scientists and bioinformaticians, will be necessary to design or update technologies and software to increase diagnostic power.

The introduction of third-generation long-read sequencing offers a promising route to advanced investigation of complex genomic variations, including balanced structural variants and chromosomal rearrangements[87]. Long-read approaches, such as single-molecule real-time sequencing and nanopore sequencing, can analyse genomes at the individual nucleotide level without conventional amplification steps and are thought to enhance mapping certainty and enable detection of mutations in repetitive DNA segments and pseudogenes[109]. Long-read sequencing is also effective at sequencing through large expanded repeats, providing a prospective tool for the study of repeat expansion-related movement disorders with superior performance in terms of accuracy and speed compared with traditional PCR-based strategies[110]. In addition, sophisticated software algorithms are being developed to support the clinical annotation of variants in non-coding DNA regions, for example, splice-disrupting intronic variants and mutations that alter regulatory functions[111].

Genomic techniques are also enabling the systematic evaluation of movement disorder-associated variants with less severe phenotypic impact and reduced penetrance. Examples include the description of *TBP* (TATA-box-binding protein) repeat expansions coexisting with pathogenic SNVs in *STUB1* (encoding E3 ubiquitin-protein ligase CHIP) in patients with ataxia, suggestive of digenic inheritance[112], and the establishment of a comprehensive database of *GBA* risk variants that contribute, with varying effect sizes, to Parkinson disease[113]. Another important development is the introduction of scalable approaches for functional mutation outcome measurements with translational potential for treatment. In the field of rare monogenic *LRRK2*-associated parkinsonism, high-throughput experimental assays have been set up to determine the biochemical consequences of any SNV identified from the genomic sequencing data set of a patient, including VUSs, paving the way for more efficient therapy trials[114].

In parallel with these advances, large-scale collaborative research initiatives are addressing the challenges of producing complete catalogues of monogenic phenotypes and increasing our understanding of how particular mutations relate mechanistically to disease biology. In the USA, the NIH Undiagnosed Diseases Network aims to improve discovery of the underlying aetiology of undiagnosed rare conditions by implementing pipelines for genomics, multiomics and functional model studies[115]. The latter studies represent an additional essential component for the characterization of unique variants and novel gene discoveries as modelling in flies, worms, zebrafish, mice, or patient-derived neuronal cells and organoids can yield unparalleled insights into the pathophysiological consequences of individual genotypic abnormalities[54]. A further complementary approach that might be used to evaluate rare gene defects in movement disorders in the future is systems biology, which can uncover mechanistically relevant genomic and multiomic signatures and overarching pathogenic drivers through network analyses and other computational methodology-based frameworks[116].

We should also continue to invest in the development of artificial intelligence-based approaches, including generative artificial intelligence, as well as corresponding standards for application in movement disorder diagnostics. Such approaches will be vital to optimizing the prioritization of different variant types, generating accurate pathogenicity predictions and enabling widespread applicability of multiomic analyses[88]. Correlation of these data with output from 'real-world' learning digital tools, such as wearable sensors, could offer additional transformative opportunities to objectively evaluate the role of certain patient-specific molecular alterations in rare movement disorders. Ongoing data-sharing activities constitute another driving force behind the scaling of clinically sound variant interpretations and additional disease-associated gene discoveries[85], and investigators should promote ethnic diversity within genomic approaches[117]. A worldwide data-sharing platform for genetic ataxias is being launched[118] and could serve as a blueprint for similar initiatives targeting other rare movement disorder subtypes. Such efforts should focus on the generalizability of knowledge for patients with heterogeneous demographic characteristics such as geographical origin, sex and age. In the context of data sharing, it will also be important to establish strategies to enhance the transfer of clinical information in the research setting, thereby facilitating bidirectional integration of insights between the clinic and the scientific arena.

Ultimately, further insights into the molecular causes of rare movement disorders will yield unique opportunities for aetiology-directed therapeutic interventions and *n*-of-1 trials by uncovering novel treatment targets or highlighting possibilities for drug re-purposing[1]. Some inspiring examples are emerging, such as the demonstration of caffeine administration as a rational effective approach to the therapy of *ADCY5*-related dyskinesia[119]. NGS-identified *ADCY5* mutations were shown to induce gain of protein function, which could be specifically reversed by adenosine A2A receptor antagonists such as the natural compound caffeine[119]. With continued progress in NGS and bioinformatic applications in rare movement disorders, we can look forward to a future in which many patients can expect a precise genetic diagnosis and, hopefully, personalized therapy.

Published online: 3 January 2024

## References


1. Rexach, J., Lee, H., Martinez-Agosto, J. A., Nemeth, A. H. & Fogel, B. L. Clinical application of next-generation sequencing to the practice of neurology. *Lancet Neurol.* **18**, 492–503 (2019).
2. Foo, J. N., Liu, J. J. & Tan, E. K. Whole-genome and whole-exome sequencing in neurological diseases. *Nat. Rev. Neurol.* **8**, 508–517 (2012).
3. Olgiati, S., Quadri, M. & Bonifati, V. Genetics of movement disorders in the next-generation sequencing era. *Mov. Disord.* **31**, 458–470 (2016).
4. Abdo, W. F., van de Warrenburg, B. P., Burn, D. J., Quinn, N. P. & Bloem, B. R. The clinical approach to movement disorders. *Nat. Rev. Neurol.* **6**, 29–37 (2010).
5. Cordeiro, D. et al. Genetic landscape of pediatric movement disorders and management implications. *Neurol. Genet.* **4**, e265 (2018).
6. Kim, M. J., Yum, M. S., Seo, G. H., Ko, T. S. & Lee, B. H. Phenotypic and genetic complexity in pediatric movement disorders. *Front. Genet.* **13**, 829558 (2022).
7. Perez-Duenas, B. et al. The genetic landscape of complex childhood-onset hyperkinetic movement disorders. *Mov. Disord.* **37**, 2197–2209 (2022).
8. Marras, C. et al. Nomenclature of genetic movement disorders: recommendations of the International Parkinson and Movement Disorder Society Task Force. *Mov. Disord.* **31**, 436–457 (2016).
9. Lange, L. M. et al. Nomenclature of genetic movement disorders: recommendations of the International Parkinson and Movement Disorder Society Task Force — an update. *Mov. Disord.* **37**, 905–935 (2022).
10. Hamosh, A., Scott, A. F., Amberger, J. S., Bocchini, C. A. & McKusick, V. A. Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. *Nucleic Acids Res.* **33**, D514–D517 (2005).
11. Gorcenco, S. et al. New generation genetic testing entering the clinic. *Parkinsonism Relat. Disord.* **73**, 72–84 (2020).
12. Kwong, A. K. et al. Exome sequencing in paediatric patients with movement disorders. *Orphanet J. Rare Dis.* **16**, 32 (2021).
13. Trinh, J. et al. Utility and implications of exome sequencing in early-onset Parkinson's disease. *Mov. Disord.* **34**, 133–137 (2019).
14. Zech, M. et al. Monogenic variants in dystonia: an exome-wide sequencing study. *Lancet Neurol.* **19**, 908–918 (2020).
15. Sun, M. et al. Targeted exome analysis identifies the genetic basis of disease in over 50% of patients with a wide range of ataxia-related phenotypes. *Genet. Med.* **21**, 195–206 (2019).

16. Martinez-Rubio, D. et al. Mutations, genes, and phenotypes related to movement disorders and ataxias. *Int. J. Mol. Sci.* **23**, 11847 (2022).
17. Landrum, M. J. et al. ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res.* **44**, D862–D868 (2016).
18. Boone, P. M., Wiszniewski, W. & Lupski, J. R. Genomic medicine and neurological disease. *Hum. Genet.* **130**, 103–121 (2011).
19. Cooper, G. M. & Shendure, J. Needles in stacks of needles: finding disease-causal variants in a wealth of genomic data. *Nat. Rev. Genet.* **12**, 628–640 (2011).
20. Pereira, R., Oliveira, J. & Sousa, M. Bioinformatics and computational tools for next-generation sequencing analysis in clinical genetics. *J. Clin. Med.* **9**, 132 (2020).
21. Karczewski, K. J. et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **581**, 434–443 (2020).
22. Crowther, L. M., Poms, M. & Plecko, B. Multiomics tools for the diagnosis and treatment of rare neurological disease. *J. Inherit. Metab. Dis.* **41**, 425–434 (2018).
23. van Karnebeek, C. D. M. et al. The role of the clinician in the multi-omics era: are you ready? *J. Inherit. Metab. Dis.* **41**, 571–582 (2018).
24. Keogh, M. J. & Chinnery, P. F. Next generation sequencing for neurological diseases: new hope or new hype? *Clin. Neurol. Neurosurg.* **115**, 948–953 (2013).
25. Coutelier, M. et al. Efficacy of exome-targeted capture sequencing to detect mutations in known cerebellar ataxia genes. *JAMA Neurol.* **75**, 591–599 (2018).
26. Keller Sarmiento, I. J. & Mencacci, N. E. Genetic dystonias: update on classification and new genetic discoveries. *Curr. Neurol. Neurosci. Rep.* **21**, 8 (2021).
27. Lange, L. M. et al. Genotype-phenotype relations for isolated dystonia genes: MDSGene systematic review. *Mov. Disord.* **36**, 1086–1103 (2021).
28. Posey, J. E. et al. Molecular diagnostic experience of whole-exome sequencing in adult patients. *Genet. Med.* **18**, 678–685 (2016).
29. Feng, H. et al. Movement disorder in GNAO1 encephalopathy associated with gain-of-function mutations. *Neurology* **89**, 762–770 (2017).
30. Wirth, T. et al. Highlighting the dystonic phenotype related to GNAO1. *Mov. Disord.* **37**, 1547–1554 (2022).
31. Turro, E. et al. Whole-genome sequencing of patients with rare diseases in a national health system. *Nature* **583**, 96–102 (2020).
32. 100,000 Genomes Project Pilot Investigators et al. 100,000 Genomes pilot on rare-disease diagnosis in health care — preliminary report. *N. Engl. J. Med.* **385**, 1868–1880 (2021).
33. Bertoli-Avella, A. M. et al. Successful application of genome sequencing in a diagnostic setting: 1007 index cases from a clinically heterogeneous cohort. *Eur. J. Hum. Genet.* **29**, 141–153 (2021).
34. Di Resta, C., Pipitone, G. B., Carrera, P. & Ferrari, M. Current scenario of the genetic testing for rare neurological disorders exploiting next generation sequencing. *Neural Regen. Res.* **16**, 475–481 (2021).
35. Pfundt, R. et al. Detection of clinically relevant copy-number variants by exome sequencing in a large cohort of genetic disorders. *Genet. Med.* **19**, 667–675 (2017).
36. Royer-Bertrand, B. et al. CNV detection from exome sequencing data in routine diagnostics of rare genetic disorders: opportunities and limitations. *Genes* **12**, 1427 (2021).
37. Zech, M. et al. Clinically relevant copy-number variants in exome sequencing data of patients with dystonia. *Parkinsonism Relat. Disord.* **84**, 129–134 (2021).
38. Coutelier, M. et al. Combining callers improves the detection of copy number variants from whole-genome sequencing. *Eur. J. Hum. Genet.* **30**, 178–186 (2022).
39. Mok, K. Y. et al. Deletions at 22q11.2 in idiopathic Parkinson's disease: a combined analysis of genome-wide association data. *Lancet Neurol.* **15**, 585–596 (2016).
40. Cunningham, A. C. et al. Movement disorder phenotypes in children with 22q11.2 deletion syndrome. *Mov. Disord.* **35**, 1272–1274 (2020).
41. Pirooznia, M., Goes, F. S. & Zandi, P. P. Whole-genome CNV analysis: advances in computational approaches. *Front. Genet.* **6**, 138 (2015).
42. Lillevali, H. et al. Genome sequencing identifies a homozygous inversion disrupting QDPR as a cause for dihydropteridine reductase deficiency. *Mol. Genet. Genom. Med.* **8**, e1154 (2020).
43. Chiang, T. et al. Atlas-CNV: a validated approach to call single-exon CNVs in the eMERGESeq gene panel. *Genet. Med.* **21**, 2135–2144 (2019).
44. Wagner, M. et al. Mitochondrial DNA mutation analysis from exome sequencing — a more holistic approach in diagnostics of suspected mitochondrial disease. *J. Inherit. Metab. Dis.* **42**, 909–917 (2019).
45. van der Sanden, B. et al. Systematic analysis of short tandem repeats in 38,095 exomes provides an additional diagnostic yield. *Genet. Med.* **23**, 1569–1573 (2021).
46. Ibanez, K. 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.* **21**, 234–245 (2022).
47. Griffin, H. R. et al. Accurate mitochondrial DNA sequencing using off-target reads provides a single test to identify pathogenic point mutations. *Genet. Med.* **16**, 962–971 (2014).
48. Poole, O. V. et al. Mitochondrial DNA analysis from exome sequencing data improves diagnostic yield in neurological diseases. *Ann. Neurol.* **89**, 1240–1247 (2021).
49. Yaldiz, B. et al. Twist exome capture allows for lower average sequence coverage in clinical exome sequencing. *Hum. Genomics* **17**, 39 (2023).
50. Dolzhenko, E. et al. ExpansionHunter Denovo: a computational method for locating known and novel repeat expansions in short-read sequencing data. *Genome Biol.* **21**, 102 (2020).
51. Rafehi, H. et al. An intronic GAA repeat expansion in FGF14 causes the autosomal-dominant adult-onset ataxia SCA50/ATX-FGF14. *Am. J. Hum. Genet.* **110**, 105–119 (2023).
52. Magrinelli, F. et al. Detection and characterization of a de novo Alu retrotransposition event causing NKX2-1-related disorder. *Mov. Disord.* **38**, 347–353 (2023).
53. Gilissen, C., Hoischen, A., Brunner, H. G. & Veltman, J. A. Disease gene identification strategies for exome sequencing. *Eur. J. Hum. Genet.* **20**, 490–497 (2012).
54. Boycott, K. M. et al. International cooperation to enable the diagnosis of all rare genetic diseases. *Am. J. Hum. Genet.* **100**, 695–705 (2017).
55. Skorvanek, M. et al. WARS2 mutations cause dopa-responsive early-onset parkinsonism and progressive myoclonus ataxia. *Parkinsonism Relat. Disord.* **94**, 54–61 (2022).
56. Sleiman, S. et al. Compound heterozygous variants in SHQ1 are associated with a spectrum of neurological features, including early-onset dystonia. *Hum. Mol. Genet.* **31**, 614–624 (2022).
57. MacDonald, J. R., Ziman, R., Yuen, R. K., Feuk, L. & Scherer, S. W. The Database of Genomic Variants: a curated collection of structural variation in the human genome. *Nucleic Acids Res.* **42**, D986–D992 (2014).
58. Lappalainen, I. et al. DbVar and DGVa: public archives for genomic structural variation. *Nucleic Acids Res.* **41**, D936–D941 (2013).
59. Brunet, T. et al. De novo variants in neurodevelopmental disorders-experiences from a tertiary care center. *Clin. Genet.* **100**, 14–28 (2021).
60. Chang, F. C. et al. Phenotypic insights into ADCY5-associated disease. *Mov. Disord.* **31**, 1033–1040 (2016).
61. Westenberger, A. et al. Spectrum of FAR1 (fatty acyl-CoA reductase 1) variants and related neurological conditions. *Mov. Disord.* **38**, 502–504 (2023).
62. Meyer, E. et al. Mutations in the histone methyltransferase gene KMT2B cause complex early-onset dystonia. *Nat. Genet.* **49**, 223–237 (2017).
63. Beetz, C. et al. LRRK2 loss-of-function variants in patients with rare diseases: no evidence for a phenotypic impact. *Mov. Disord.* **36**, 1029–1031 (2021).
64. Blauwendraat, C. et al. Frequency of loss of function variants in LRRK2 in Parkinson disease. *JAMA Neurol.* **75**, 1416–1422 (2018).
65. Salles, P. A., Mata, I. F., Brunger, T., Lal, D. & Fernandez, H. H. ATP1A3-related disorders: an ever-expanding clinical spectrum. *Front. Neurol.* **12**, 637890 (2021).
66. Havrilla, J. M., Pedersen, B. S., Layer, R. M. & Quinlan, A. R. A map of constrained coding regions in the human genome. *Nat. Genet.* **51**, 88–95 (2019).
67. Wiel, L. et al. MetaDome: pathogenicity analysis of genetic variants through aggregation of homologous human protein domains. *Hum. Mutat.* **40**, 1030–1038 (2019).
68. Wiel, L. et al. De novo mutation hotspots in homologous protein domains identify function-altering mutations in neurodevelopmental disorders. *Am. J. Hum. Genet.* **110**, 92–104 (2023).
69. Singh, S. et al. De novo variants of NR4A2 are associated with neurodevelopmental disorder and epilepsy. *Genet. Med.* **22**, 1413–1417 (2020).
70. Winter, B. et al. NR4A2 and dystonia with dopa responsiveness. *Mov. Disord.* **36**, 2203–2204 (2021).
71. Jesus, S. et al. NR4A2 mutations can cause intellectual disability and language impairment with persistent dystonia-parkinsonism. *Neurol. Genet.* **7**, e543 (2021).
72. Geisheker, M. R. et al. Hotspots of missense mutation identify neurodevelopmental disorder genes and functional domains. *Nat. Neurosci.* **20**, 1043–1051 (2017).
73. Kircher, M. et al. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.* **46**, 310–315 (2014).
74. Wilcox, E. H. et al. Evaluating the impact of in silico predictors on clinical variant classification. *Genet. Med.* **24**, 924–930 (2022).
75. Martin, A. R. et al. PanelApp crowdsources expert knowledge to establish consensus diagnostic gene panels. *Nat. Genet.* **51**, 1560–1565 (2019).
76. Richards, 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.* **17**, 405–424 (2015).
77. Li, Q. & Wang, K. InterVar: clinical interpretation of genetic variants by the 2015 ACMG-AMP guidelines. *Am. J. Hum. Genet.* **100**, 267–280 (2017).
78. Vears, D. F., Senecal, K. & Borry, P. Reporting practices for variants of uncertain significance from next generation sequencing technologies. *Eur. J. Med. Genet.* **60**, 553–558 (2017).
79. Clift, K. et al. Patients' views on variants of uncertain significance across indications. *J. Community Genet.* **11**, 139–145 (2020).
80. Liu, P. et al. Reanalysis of clinical exome sequencing data. *N. Engl. J. Med.* **380**, 2478–2480 (2019).
81. Mensah, N. E. et al. Automated reanalysis application to assist in detecting novel gene-disease associations after genome sequencing. *Genet. Med.* **24**, 811–820 (2022).
82. Fattahi, Z. et al. Iranome: a catalog of genomic variations in the Iranian population. *Hum. Mutat.* **40**, 1968–1984 (2019).
83. Cordts, I. et al. Adult-onset neurodegeneration in nucleotide excision repair disorders (NERD(ND)): time to move beyond the skin. *Mov. Disord.* **37**, 1707–1718 (2022).
84. Skorvanek, M. et al. Adult-onset neurodegeneration in nucleotide excision repair disorders: more common than expected. *Mov. Disord.* **37**, 2323–2324 (2022).
85. Rehm, H. L. Evolving health care through personal genomics. *Nat. Rev. Genet.* **18**, 259–267 (2017).
86. Gilissen, C., Hoischen, A., Brunner, H. G. & Veltman, J. A. Unlocking Mendelian disease using exome sequencing. *Genome Biol.* **12**, 228 (2011).

87. Bamshad, M. J., Nickerson, D. A. & Chong, J. X. Mendelian gene discovery: fast and furious with no end in sight. *Am. J. Hum. Genet.* **105**, 448–455 (2019).
88. McInnes, G. et al. Opportunities and challenges for the computational interpretation of rare variation in clinically important genes. *Am. J. Hum. Genet.* **108**, 535–548 (2021).
89. Stenson, P. D. 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.* **136**, 665–677 (2017).
90. Firth, H. V. et al. DECIPHER: database of chromosomal imbalance and phenotype in humans using Ensembl resources. *Am. J. Hum. Genet.* **84**, 524–533 (2009).
91. Brandon, M. C. et al. MITOMAP: a human mitochondrial genome database — 2004 update. *Nucleic Acids Res.* **33**, D611–D613 (2005).
92. Lill, C. M. et al. Launching the movement disorders society genetic mutation database (MDSGene). *Mov. Disord.* **31**, 607–609 (2016).
93. Hindorff, L. A. et al. Prioritizing diversity in human genomics research. *Nat. Rev. Genet.* **19**, 175–185 (2018).
94. Philippakis, A. A. et al. The matchmaker exchange: a platform for rare disease gene discovery. *Hum. Mutat.* **36**, 915–921 (2015).
95. Sobreira, N., Schiettecatte, F., Valle, D. & Hamosh, A. GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. *Hum. Mutat.* **36**, 928–930 (2015).
96. Gannamani, R., van der Veen, S., van Egmond, M., de Koning, T. J. & Tijssen, M. A. J. Challenges in clinicogenetic correlations: one phenotype — many genes. *Mov. Disord. Clin. Pract.* **8**, 311–321 (2021).
97. Neilson, D. E. et al. A novel variant of ATP5MC3 associated with both dystonia and spastic paraplegia. *Mov. Disord.* **37**, 375–383 (2022).
98. Turner, T. N. et al. denovo-db: a compendium of human de novo variants. *Nucleic Acids Res.* **45**, D804–D811 (2017).
99. Lappalainen, I. et al. The European genome-phenome archive of human data consented for biomedical research. *Nat. Genet.* **47**, 692–695 (2015).
100. Zurek, B. et al. Solve-RD: systematic pan-European data sharing and collaborative analysis to solve rare diseases. *Eur. J. Hum. Genet.* **29**, 1325–1331 (2021).
101. Zech, M. et al. Variants in mitochondrial ATP synthase cause variable neurologic phenotypes. *Ann. Neurol.* **91**, 225–237 (2022).
102. Aref-Eshghi, E. et al. Evaluation of DNA methylation episignatures for diagnosis and phenotype correlations in 42 Mendelian neurodevelopmental disorders. *Am. J. Hum. Genet.* **106**, 356–370 (2020).
103. Mirza-Schreiber, N. et al. Blood DNA methylation provides an accurate biomarker of KMT2B-related dystonia and predicts onset. *Brain* **145**, 644–654 (2022).
104. Ciolfi, A. et al. Childhood-onset dystonia-causing KMT2B variants result in a distinctive genomic hypermethylation profile. *Clin. Epigenetics* **13**, 157 (2021).
105. Kremer, L. S. et al. Genetic diagnosis of Mendelian disorders via RNA sequencing. *Nat. Commun.* **8**, 15824 (2017).
106. Lee, H. et al. Diagnostic utility of transcriptome sequencing for rare Mendelian diseases. *Genet. Med.* **22**, 490–499 (2020).
107. Yepez, V. A. et al. Clinical implementation of RNA sequencing for Mendelian disease diagnostics. *Genome Med.* **14**, 38 (2022).
108. Fresard, L. et al. Identification of rare-disease genes using blood transcriptome sequencing and large control cohorts. *Nat. Med.* **25**, 911–919 (2019).
109. Amarasinghe, S. L. et al. Opportunities and challenges in long-read sequencing data analysis. *Genome Biol.* **21**, 30 (2020).
110. Miyatake, S. et al. Rapid and comprehensive diagnostic method for repeat expansion diseases using nanopore sequencing. *NPJ Genom. Med.* **7**, 62 (2022).
111. Wagner, N. et al. Aberrant splicing prediction across human tissues. *Nat. Genet.* **55**, 861–870 (2023).
112. Magri, S. et al. Digenic inheritance of STUB1 variants and TBP polyglutamine expansions explains the incomplete penetrance of SCA17 and SCA48. *Genet. Med.* **24**, 29–40 (2022).
113. Parlar, S. C., Grenn, F. P., Kim, J. J., Baluwendraat, C. & Gan-Or, Z. Classification of GBA1 variants in Parkinson's disease: the GBA1-PD browser. *Mov. Disord.* **38**, 489–495 (2023).
114. Kalogeropulou, A. F. et al. Impact of 100 LRRK2 variants linked to Parkinson's disease on kinase activity and microtubule binding. *Biochem. J.* **479**, 1759–1783 (2022).
115. Splinter, K. et al. Effect of genetic diagnosis on patients with previously undiagnosed disease. *N. Engl. J. Med.* **379**, 2131–2139 (2018).
116. Buphamalai, P., Kokotovic, T., Nagy, V. & Menche, J. Network analysis reveals rare disease signatures across multiple levels of biological organization. *Nat. Commun.* **12**, 6306 (2021).
117. Bakhit, Y. et al. Intrafamilial and interfamilial heterogeneity of PINK1-associated Parkinson's disease in Sudan. *Parkinsonism Relat. Disord.* **111**, 105401 (2023).
118. Beijer, D. et al. Standards of NGS data sharing and analysis in ataxias: recommendations by the NGS working group of the Ataxia Global Initiative. *Cerebellum* https://doi.org/10.1007/s12311-023-01537-1 (2023).
119. Meneret, A. et al. Efficacy of caffeine in ADCY5-related dyskinesia: a retrospective study. *Mov. Disord.* **37**, 1294–1298 (2022).
120. Gilbert, D. L., Leslie, E. J., Keddache, M. & Leslie, N. D. A novel hereditary spastic paraplegia with dystonia linked to chromosome 2q24-2q31. *Mov. Disord.* **24**, 364–370 (2009).



## Acknowledgements
M.Z. and J.W. receive research support from the German Research Foundation (DFG 458949627; ZE 1213/2-1; WI 1820/14-1). M.Z. acknowledges grant support from the European Joint Programme on Rare Diseases (European Joint Programme on Rare Diseases 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 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. M.Z. is a member of the Medical and Scientific Advisory Council of the Dystonia Medical Research Foundation and a member of the Governance Council of the International Cerebral Palsy Genomics Consortium. MZ's research is supported by a "Schlüsselprojekt" grant from the Else Kröner-Fresenius-Stiftung (2022_EKSE.185).


## Author contributions
Both authors designed and supervised the work and wrote the article.

## Competing interests
The authors declare no competing interests.

## Additional information
**Peer review information** *Nature Reviews Neurology* thanks H. Houlden and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



## Related links
**ClinVar:** https://www.ncbi.nlm.nih.gov/clinvar/
**CNVnator:** https://github.com/abyzovlab/CNVnator
**Database of Genomic Variants:** http://dgv.tcag.ca/dgv/app/home
**dbVar:** https://www.ncbi.nlm.nih.gov/dbvar/
**DECIPHER:** https://www.deciphergenomics.org/
**DELLY:** https://github.com/dellytools/delly
**European Genome–Phenome Archive:** https://web2.ega-archive.org/
**European Joint Programme on Rare Diseases:** https://www.ejprarediseases.org/
**ExomeDepth:** https://cran.r-project.org/package=ExomeDepth
**ExpansionHunter Denovo:** https://github.com/Illumina/ExpansionHunterDenovo
**GeneMatcher:** https://genematcher.org/
**Genome Aggregation Database (gnomAD):** https://gnomad.broadinstitute.org/
**Genome Analysis Toolkit:** https://gatk.broadinstitute.org/hc/en-us
**German Human Genome–Phenome Archive:** https://www.ghga.de/
**Human Gene Mutation Database:** https://www.hgmd.cf.ac.uk/ac/index.php
**InterVar:** https://wintervar.wglab.org/
**Manta:** https://github.com/Illumina/manta
**Matchmaker Exchange (MME):** https://www.matchmakerexchange.org/
**MetaDome:** https://stuart.radboudumc.nl/metadome/
**MITOMAP:** https://www.mitomap.org/MITOMAP
**Movement Disorder Society Genetic mutation database:** https://www.mdsgene.org/
**NIH Undiagnosed Diseases Network:** https://commonfund.nih.gov/Diseases
**Online Mendelian Inheritance in Man:** https://www.omim.org/
**PanelApp:** https://panelapp.genomicsengland.co.uk/
**Solve-RD:** https://solve-rd.eu/