Parkinsonism and Related Disorders 134 (2025) 107311

Contents lists available at ScienceDirect

Parkinsonism and Related Disorders

journal homepage: www.elsevier.com/locate/parkreldis

Big data and transformative bioinformatics in genomic diagnostics and beyond

Alice Saparov [a,b,c], Michael Zech [a,b,c,*]

[a] Institute of Human Genetics, Technical University of Munich, School of Medicine and Health, Munich, Germany
[b] Institute of Neurogenomics, Helmholtz Munich, Neuherberg, Germany
[c] Institute for Advanced Study, Technical University of Munich, Garching, Germany

ARTICLE INFO

Keywords:
big data
Bioinformatics
Phenomics
Genomics
Multi-omics
Dystonia

ABSTRACT

The current era of high-throughput analysis-driven research offers invaluable insights into disease etiologies, accurate diagnostics, pathogenesis, and personalized therapy. In the field of movement disorders, investigators are facing an increasing growth in the volume of produced patient-derived datasets, providing substantial opportunities for precision medicine approaches based on extensive information accessibility and advanced annotation practices. Integrating data from multiple sources, including phenomics, genomics, and multi-omics, is crucial for comprehensively understanding different types of movement disorders. Here, we explore formats and analytics of big data generated for patients with movement disorders, including strategies to meaningfully share the data for optimized patient benefit. We review computational methods that are essential to accelerate the process of evaluating the increasing amounts of specialized data collected. Based on concrete examples, we highlight how bioinformatic approaches facilitate the translation of multidimensional biological information into clinically relevant knowledge. Moreover, we outline the feasibility of computer-aided therapeutic target evaluation, and we discuss the importance of expanding the focus of big data research to understudied phenotypes such as dystonia.

1. Introduction

The availability of diagnostically relevant patient-derived data in diverse clinical fields, including neurological movement disorders, has increased tremendously because of major advances in laboratory techniques and scalable, high-performance computing [1]. Monogenic causes are thought to account for most cases of rare movement disorders [2]. Moreover, it is now understood that rare genetic variation contributes to disease susceptibility in a substantial proportion of movement syndromes that are more commonly seen in daily practice, such as Parkinson disease (PD) and late-onset ataxia (LOA) [3,4]. Efforts to integrate data-intensive analysis strategies, such as genomic sequencing and multi-omics approaches, are thus expanding worldwide [5,6]. In recent years, decreasing costs of massively parallel assays and the creation of novel software packages and bioinformatics workflows have led to an accumulation of data from both research-based studies and clinical investigations [7,8]. The landscape of high-throughput analytic applications, encompassing genomics, epigenomics, transcriptomics, proteomics, and metabolomics, has drawn growing attention for use in diagnostics, exploration of pathogenesis, screening of drug targets, and the creation of predictive models of disease in asymptomatic populations [2,9]. Over 15 years ago, the scientific community focused on the identification of genetic variants. Now, the emphasis has transitioned to data interpretation, integration of multiple layers of information, development of optimized computational processing pipelines, and the sharing of data internationally [10,11]. Ideally, the output from different big data-generating methods and tools would be combined in a standardized manner to enable holistic applications for precise diagnosis, disease classification, biomarker and therapeutic target discovery, as well as prediction of prognosis, with the ultimate goal of delivering individualized care for each patient [12]. Despite the many challenges ahead, we are witnessing a confluence of breakthroughs in wet lab methodologies, bioinformatics systems, innovative research projects, and clinical cooperative initiatives, supporting the realization of big-data science toward precision medicine [13,14]. Herein, we present opportunities derived from our increasing ability to leverage large-scale

* Corresponding author. Institute of Neurogenomics, Helmholtz Zentrum München, Deutsches Forschungszentrum für Gesundheit und Umwelt (GmbH), Ingolstädter Landstraße 1, Neuherberg, 85764, Germany.
E-mail address: michael.zech@mri.tum.de (M. Zech).

https://doi.org/10.1016/j.parkreldis.2025.107311
Received 11 December 2024; Received in revised form 23 January 2025; Accepted 25 January 2025
Available online 3 February 2025

phenotype-driven datasets, summarizing key genomic technologies, associated bioinformatic methods, including artificial intelligence (AI)-based algorithms, and their roles for deriving useful insights into disease etiologies. As big data grows, there is a need to adopt integrative pathways for parallel analysis of multiple omics and global cooperation to facilitate efficient use and re-use of the mass amounts of generated information [15–17]. Finally, we outline the prospects of computational platforms that can support treatment-target identification, and we stress future directions that should be considered to improve big-data analytics for understudied phenotypes. Although the discussed approaches are of wide relevance, we focus on their application in movement disorders, providing examples from current movement-disorder literature and ongoing programs designed to advance research on dystonic, Parkinsonian, and/or ataxic/-spastic conditions (Figs. 1 and 2).

1.1. Phenomics

Comprehensive phenotyping and systematic phenotypic analyses represent crucial prerequisites for harnessing the maximum benefit of biomedical big data [2,18,19]. Considering an estimated number of more than 50 million individuals enrolled in clinical cohorts and biobanks worldwide [20], the cataloging of these subjects` trait manifestations is a huge task and requires advanced computational processing and bioinformatics infrastructure. Efforts collecting phenotypic data from thousands of patients with dystonia [21,22], PD [3], ataxia [23], and other movement disorders [2] are underway, motivated by the goal of expanding the recruitment to diverse underrepresented populations [24,25]. So far, we have only begun to tackle the large amounts of data produced by deep (er) phenotyping of movement disorder-affected patients, which are often captured by not only one physician but rather a team of multidisciplinary experts involved throughout the course of a differential diagnosis and a patient's care over time [26]. One way we can track the state of our accumulated knowledge about relationships between phenotypes and their associated genes is through publicly available online databases such as Online Mendelian Inheritance in Man (OMIM) [27]. While around 5000 genes are presently recognized as clinically relevant [20], an OMIM-based query for genes linked to dystonia/tremor, parkinsonism, and ataxia yields 598, 124, and 733 hits (accessed 12/2024) [27], respectively, highlighting heterogeneity as an important source of data intensity. Despite the rich repository of evidence, OMIM is curated manually, making it difficult to ensure all entries are up to date. A recent study investigating a compilation of aggregate phenotypic data from over 50,000 patients with neurodevelopmental disorders suggested that up to one-third of disease-specific clinical features may be under-documented in OMIM [28]. Unfortunately, the descriptions of movement abnormalities seem to be particularly often incomplete [2,29]. For example, Rots et al. [30] reported that 24 % of patients with pathogenic *KDM6B* variants displayed spastic-ataxic symptoms or dystonia, although these features had never been reported for *KDM6B*-related disease in OMIM, highlighting challenges of using open-source dataset catalogs for diagnostic evaluation in movement disorders. Alternative analysis platforms designed to accelerate the interactive exploration of genotype-phenotype correlations are now under development, often supported by computational technology that is capable of optimizing phenotype extraction from multiple sources [31]. Integration of AI-based text mining tools can narrow relevant information from high-throughput phenotyping and provide guidance for analysts to establish precise diagnoses [32,33]. Two recent studies demonstrated that natural language processing systems, automatically extracting phenomics data from electronic health records of patients with heterogeneous rare-disease symptoms, including abnormal movements, were powerful tools to aid clinicians in augmenting diagnostic certainty [34,35]. When combined with genomic data, an electronic health record-mining platform was able to quickly flag causative pathogenic variants with no need for expert diagnosticians [35]. As more multimodal big-data analyses will be carried

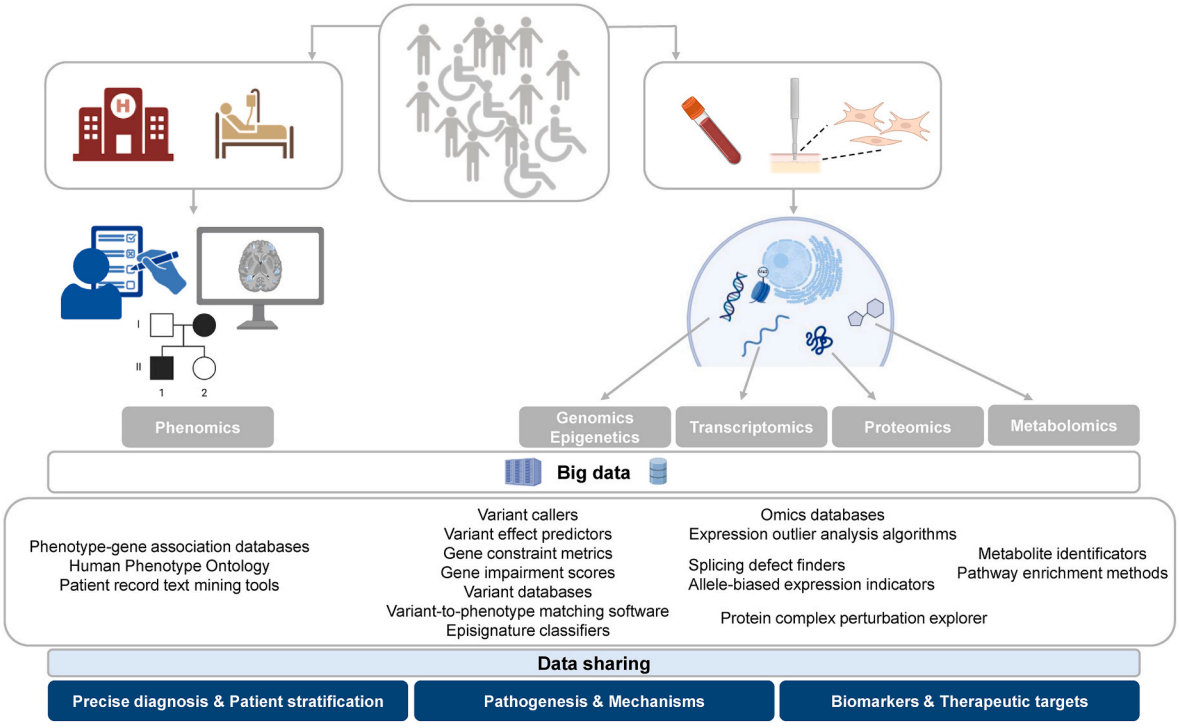

Fig. 1. Multi-omic information contributing to big data and bioinformatic strategies for data analysis.
Phenomics data are collected in the clinical setting. Patient biosamples, including blood and/or skin biopsy-derived primary fibroblasts, are used to obtain genomics, epigenetics, transcriptomics, proteomics, and metabolomics data. The outputs of the different single omic technologies can be combined to provide a greater depth to analysis and to enhance research and diagnostic discoveries. The ultimate goal of multi-omics big data studies is to advance precision clinical care with improved diagnosis, understanding of pathomechanisms, as well as identification of biomarkers and efficacious treatment options.

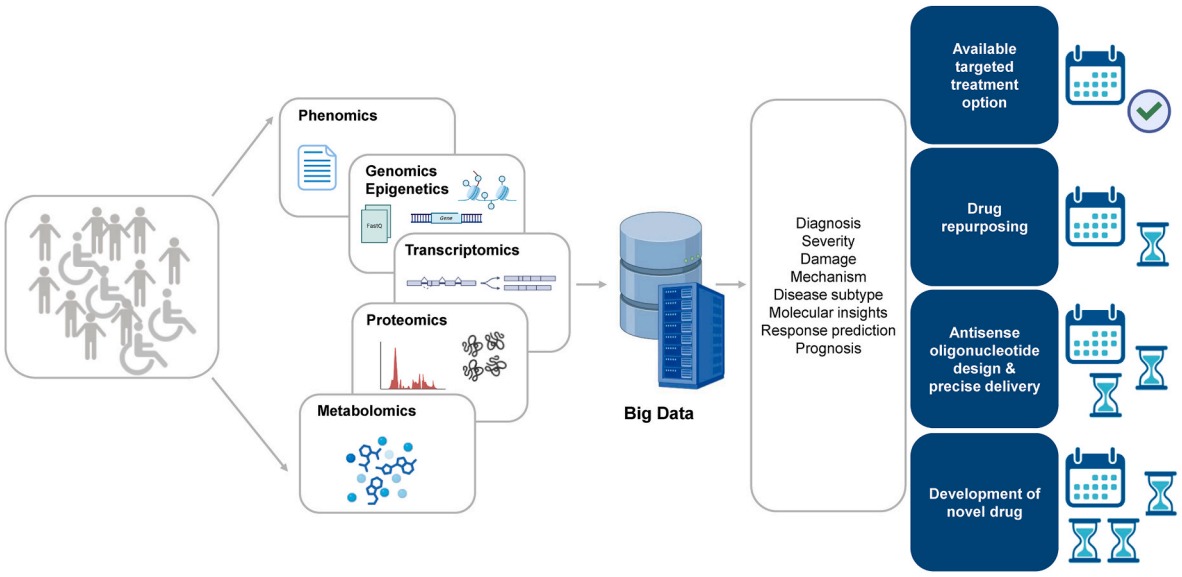

Fig. 2. Proposed approach for big data translation into therapy.
Following bioinformatic integration of multi-omics data, the results may drive insights into targeted interventions for individual patients or larger groups of etiologically stratified patients. The development of novel drugs is time consuming, whereas the application of existing treatments based on specific pathogenic mechanisms can be straightforward.

out for disease characterization, frameworks are necessary to allow for standardized capture of phenotypes for computational analysis [6,36,37]; the Human Phenotype Ontology (HPO) provides computable descriptions of clinical features, listing, for example, 19 terms that define different manifestations of dystonia [38]. Although HPO may be regarded as an imperfect means of reporting motor phenomenology, this vocabulary has been effectively incorporated into large-scale genomic analysis projects for patients with dystonia [39] and other movement disorders [37]. Lastly, studies showed that AI-aided algorithms trained on clinical big data can outperform human specialists in distinguishing between specific disease states through pattern recognition [40,41]; such progress may also transform movement-disorder care when digital high-resolution phenotyping with sensors and other still under-utilized monitoring systems will enter the clinic [42].

1.2. Genomics

Advancements in next-generation sequencing technologies have spurred the production of large volumes of biomedical information [2,43]. Daily implementation of high-throughput methods that generate data at the genome-scale enables comprehensive interrogation of genetic aberrations underlying phenotypic presentations in the movement-disorder clinic [2]. In this process, bioinformatics offers scalable solutions for analyzing the data-intensive output from whole-exome sequencing (WES) and whole-genome sequencing (WGS) [44]. Over the past years, modern WES/WGS-data pipelines coupled with multifunctional software tools facilitated a more accurate understanding of the etiological spectrum of movement phenotypes. For example, recent seminal studies elucidated the genomic architecture of cerebral palsy by cohort-wide bioinformatics-driven whole-genome annotation [45,46], serving as an entry point to precision care in this underserved field [46]. Moreover, big genomic data analyses provided insights into the role of phenotypic predictors of diagnostic success in movement disorders (e.g., early-onset generalized dystonia with neurodevelopmental comorbidity was shown to be associated with a molecular yield of up to ~50 %) [21,47]. A wide range of computer-aided methodological approaches is necessary to diagnose patients by WES or WGS [48,49], thus, for the identification of different types of variations in sequencing raw data, a suite of state-of-the-art calling algorithms has been designed [50]. In addition to widely used conventional tools, including the Genome Analysis Toolkit [51], AI-based callers such as Google`s DeepVariant are now available and may supersede other workflows for discovering single-nucleotide variants and short insertions/deletions (INDELs) [52]. Additionally, a collection of copy-number variant-detection tools with different underlying models and input-data requirements allows analysts to efficiently investigate larger duplications and deletions aligned against a reference genome [53]. The diagnostic evaluation of some movement disorders, especially ataxia, has been revolutionized by the introduction of automated tools that enable the discovery of expanded short-tandem repeats (STRs) in WES/WGS data [54]. ExpansionHunter addresses long-standing needs to characterize clinically relevant STR variation in a fast and robust manner [55]; this framework was essential to tackle mystery ataxic disorders that had remained undiagnosed in expert settings for decades (e.g., *FGF14*-related forms of LOA) [56]. Finally, the recent advent of long-read sequencing (LRS) platforms such as Oxford Nanopore technology (ONT) and Pacific Biosciences leads to further improvements in variant-identification accuracy, although these methods come with challenging demands for new, computationally intensive data processing strategies, which are not yet widely available across laboratories [57,58]. Nevertheless, pilot studies using LRS in patients with movement disorders highlighted insightful results, such as the demonstration of superior structural-variant detection by ONT in *PARK2*-related PD [59]. One of the most critical stages of WES/WGS-based diagnostics is data interpretation, which relies heavily on inference of the impact of called variants [2]. Although we still have a paucity of knowledge to interpret many (if not most) rare variants that are found in a patient exome or genome, the integration of different types of evidence, including bioinformatic gene-prioritizations and predictions of variant effects based on computational tools, can provide essential clues for diagnosis [2]. Today, high-quality annotations of mutational constraint and expected deleteriousness are available from multiple models. Karczewski et al. used data from the Genome Aggregation Database (gnomAD) to define gold-standard gene-level intolerance scores, providing a reliable means of improving the interpretation of variants [60]. The observed-to-expected upper bound fraction scores from gnomAD represent computationally derived ranking metrics, helping investigators prioritize genes in which variants are likely of clinical

significance [61]. In a recent WES/WGS study of a cohort of patients with neurodevelopmental movement disorders, Poggio et al. found missense variants in *ATP2B2*, prioritized based on the gene`s significant constraint against missense substitutions [62]; follow-up functional experiments validated the disease-causing nature of the variants [62], highlighting how metrics calculated from harmonized large-scale sequencing data can inform on disease-association hypotheses per gene. Moreover, innovative developments of computational variant-effect predictors have begun to transform automated analysis pipelines for sequencing data [63]. Key drivers of improved predictor effectiveness have been the integration of advanced machine-learning-based approaches and the use of meta-predictors, merging the output from several existing individual tools [63]. Google DeepMind`s AlphaMissense, an unsupervised language model classifier, is capable of providing proteome-wide structural predictions for all possible amino-acid substitutions and has been proposed to achieve at least 90 % precision when predicting the established clinical impact of missense mutations [64]. Other AI-based tools like PrimateAI-3D from Illumina have been trained on extensive multiple-sequence alignments and augment variant predictions with cross-species information [65]. INDELs are the focus of some new predictive models, such as INDELpred [66], aiming to improve the interpretation of this common type of genomic variation for accurate diagnoses. Computational prediction-based variant annotations are also increasingly relevant for the clinically oriented analysis of splice sites and non-coding parts of the genome [67]. As more genomes of patients with movement disorders are sequenced, a deeper understanding of DNA regions that do not correspond to protein-coding sequences is also important. SpliceAI is a deep-learning splicing prediction algorithm allowing for faster assessment of putative splice defects [68], which represent ~10 % of all disease-causing mutations and are thought to explain a notable fraction of monogenic movement disorders [69]. Despite additional efforts concentrating on the definition of mutation-intolerant non-coding sequences with deep learning [70], the interpretability of non-coding variants remains a big challenge; we would like to encourage the movement-disorder community to invest in innovative computational science to overcome existing barriers in the analysis of pathogenic variants in non-coding elements, e.g., by providing the datasets of their patients as potential new AI use-cases. The precise constellation by which all the different computational methods and algorithms for variant analysis could be combined into an optimal pipeline is not fully streamlined, but most laboratories guarantee diagnostic performance by integrating cross-validated sets of interconnected bioinformatic systems [21,48,49]. Ultimately, dynamic adaption mechanisms in software architecture are recommended, given the surge of newly emerging machine learning, AI, and natural language-processing applications in genomics [11].

1.3. Multi-omics

There is a growing need to improve diagnostic success rates by performing combined analyses of multi-omics data produced through complementary technologies [2,6,9,71]. Simultaneously, technological developments that deal with the validation and scalability of the different big-data formats are required [72]. One recent major advancement in the realization of multi-level molecular profiling in patients affected by rare diseases, including movement disorders, has been the introduction of clinical DNA methylation episignature testing [73]. This methodology allows the determination of characteristic patterns in the methylomes of individual patients, which are disease-specific and can be used to (re-)classify genomic variant findings [73]. Most currently established episignatures have been constructed on the basis of microarray technology-based standardized, high-throughput assessments of methylation profiles in peripheral blood [74]. In dystonia caused by variants in the histone-modifier gene *KMT2B*, episignature studies were useful to predict disease subtypes (e.g., early-versus later-onset forms) [75]; moreover, variant effect-size analyses based on methylation changes in larger sets of patients suggested that *KMT2B* variants define allelic series, offering potential insights into mechanisms of variable expressivity [76]. The increasing application of other high-throughput screening methodologies, such as transcriptomics and proteomics, can also allow for in-depth evaluations of variant effects on molecular and clinical phenotypes [77,78]. These technologies enable quantitative and qualitative assessments of all the RNAs and proteins that are detectable in an investigated cell type, capable of informing on the consequences of regulatory and cryptic splicing variants [2,9,79]. Murdock et al. [80] provided a global perspective on transcriptome alterations in fibroblasts from patients with rare undiagnosed diseases, identifying a movement disorder-affected patient with an RNA expression defect-inducing promoter variant in *SPR*. Moreover, proteomic analyses are now implemented in large-scale genomics studies to yield molecular perturbations underlying disease [6,77]. Our group integrated fibroblast-derived proteomic data combined with genomic and transcriptomic information at scale to prioritize causal variants in long-term unresolved cases with dystonic syndromes; our results highlighted the discovery power of high-depth proteomics, showing that the approach helped to uncover driver genes with difficult-to-detect non--coding mutations (Zech et al., *in preparation*). Sophisticated denoising autoencoder-based methods have been developed for reproducible identification of expression outliers in patient transcriptomes and proteomes [77,81], along with complementary processing tools that support unbiased searches for splice defects with high sensitivity and specificity [82]. Although big-data multi-omic analyses are currently limited by the lack of expression of several neurologically relevant genes in easily accessible tissue, future strategies such as comprehensive CRISPR gene-activation may expand our possibilities to comprehensively apply RNA/proteome diagnostics [83]. Furthermore, metabolomics, focusing on the interrogation of large sets of small molecules representative of core biological processes, is gaining momentum for the diagnosis of patients with movement disorders, especially PD [84]. Metabolomic approaches could be routinely used to screen for global pathway abnormalities in patients versus controls, highlighting disease-specific biomarkers that would enable stratification for tailored treatment [85]. Many details remain to be addressed surrounding the widespread implementation of multi-omics diagnostic procedures in the field of movement disorders. We will need guidance on how to combine the results generated on different platforms with heterogeneous precision levels to obtain optimal insights across molecular layers with valid interpretability. We are hopeful that the systematic expansion of samples and the development of even more efficient processing software based on AI will provide clinicians with improved views of their patients` multi-omic disease profiles.

1.4. Data sharing

Every individual is expected to carry over 200 very rare variants in the protein-coding DNA sequence [60], and the majority of variants reported in patients with rare disorders are unique [69]. Indeed, studies showed that the number of distinct causally involved genotypes increases linearly with the number of investigated disease subjects [21,47,86]. It is well established that we need to assemble much larger datasets to enhance the effectiveness and speed of diagnosis and to reach saturation in the catalogs of clinically relevant variants [20]. In recent years, scalable databases and controlled-access methods capable of dealing with the tremendous amounts of shareable genomic data have begun to contribute to a better understanding of disease causation [2], all while meeting requirements to maintain the patients` privacy [87]. Variant-level data are increasingly archived in community-curated repositories such as ClinVar [69], HGMD [88], and the movement disorder-specific database MDSGene [89]. Importantly, these resources facilitate the connection between responsible analysts, accelerating insights into unappreciated genetic heterogeneity. For example, our group

recently identified a patient with dystonia and congenital eye malformation who harbored a nonsense variant in *RARB*, a gene for which only missense mutations were described [90]; the cross-referencing to ClinVar highlighted that the exact same nonsense change had been found in two independent patients from other laboratories, confirming a new pathogenic mechanism and securing the diagnosis [90]. ClinVar data have also been used in a very recent large-scale case-control meta-analysis of rare variants linked to PD [91]. Based on evaluation of genotypes from over 3 million individuals, the researchers showed that variants in PD-related genes that are deposited in ClinVar could be categorized into high-confidence and low-confidence PD-associated variants [91]; the findings can provide an improved understanding of the clinical significance of ClinVar-listed variants for PD at scale, guiding genetic testing design and follow-up studies into genetic architecture. Additionally, data sharing via the MatchMaker exchange platform [92] including GeneMatcher [93] has been instrumental to advance knowledge about genotype-phenotype associations in the past decade [15]. GeneMatcher-supported multicenter analyses led to an unprecedented acceleration in the identification of individuals affected by the same rare condition [15], such as *CUL3*-related neurodevelopmental disease, which has been recently demonstrated to be linked to a phenotypic spectrum ranging from unspecific developmental delay in infancy to specific movement disorders in adulthood [94]. To maximize patient benefit through reusable data, the output from WES, WGS, and multi-omics studies, as well as standardized phenotypic descriptors, can be registered in environments enabling cross-cohort investigations such as the European Genome-Phenome Archive and its national hubs, e.g., the German Human Genome-Phenome Archive [95]. Movement disorder-centered initiatives for collecting and sharing multilayered data have been set up, e.g., for ataxia [23,96] and PD [3]; the GP2 consortium focuses on the inclusion of data from geographically underrepresented PD populations [97]. In addition to the sharing of data, our capabilities for collaborative analysis and joint interpretation must be catalyzed. Here, international "solvathons" such as those organized by the Solve-RD network [98] may prove useful, since these events foster data-centric interactions between researchers, clinicians, bioinformaticians, and other experts [99]. New activities are on the horizon that will further enhance our endeavors of boosting big-data analyses for movement disorders on a global scale.

1.5. Beyond diagnostics - translating big data to treatment

Recent technical and computational developments enabling the evaluation of high-throughput, large-scale phenomics, genomics, and multi-omics data will be instrumental in advancing strategies for the realization of precision medicine [5] (Fig. 2). Beyond the improvement of diagnostics, big-data analyses and related research can offer unique pathways to the translation of clinically relevant information from multiple layers into personalized treatment regimens [100]. Innovative approaches integrating big-data results with clinical interventions and targeted therapies are emerging [100]. For example, Lewis et al. [101] leveraged seminal advances in the understanding of the genetic architecture of cerebral palsy to demonstrate that 8 % of molecular findings in 1841 WES-diagnosed individuals enabled precision medicine interventions; this knowledgebase for genomics-guided patient benefit is expected to greatly aid the introduction of optimized management workflows targeting the mechanisms of disease in the field of cerebral palsy, a group of conditions which had traditionally been considered untreatable [102]. Moreover, genomics/multi-omics insights can be helpful to promote drug repurposing: a growing body of evidence suggests that disease-associated genetic variation or molecular connectome alterations hinting at specific etiologic targets or perturbed signatures may provide substantial opportunities to use existing medications in novel clinical applications [103,104], such as movement disorders. Notably, among 50 drugs approved by the FDA within one year, up to two-thirds are reported to have genetically-informed bases of action [105]. In PD, for example, big-data inference approaches have begun to predict repurposing candidates based on genome-wide association study signals and biological pathway analyses [106]. Although cost-consuming and inappropriate for fast-track clinical implementation, the design of novel compounds suitable for tailored therapy is also possible once specific genomically-guided intervention targets or druggable movement disorder-driving proteins have been identified [104]; in such process, bioinformatics and AI-empowered strategies will play key roles, building models to predict a given target`s druggability and to introduce generative strategies that can use synthetic data for advanced target selection [103]. Finally, progress has also been made in the translation of genomic findings into testing of more specific treatments based on mutational mechanisms. Mittal et al. [107] employed big-data mining to characterize the landscape of monogenic variants treatable with antisense-oligonucleotide (ASO) therapy: by computationally assessing ClinVar information they predicted that roughly half of all annotated pathogenic variants would have a promising avenue for ASOs, acting against loss-of-function, gain-of-function and/or splice defect-specific effects. Mechanism-centric ASO approaches are already under development for movement disorders, including ataxia telangiectasia [108], and could be generalized to other indications. Further advances in deep learning predictions given the abundance of data from patients with movement disorders are likely to reshape the process of screening for new drugs and their implementation into clinical practice.

2. Perspectives

With the accumulation of new datasets and samples from patients with movement disorders worldwide, the role of big-data research is becoming increasingly important. We anticipate that frameworks for computationally aided mechanism and drug-repurposing explorations based on a combined phenomics-genomics-multiomics concept will expand to markedly pursue the improvement of patient care in the near future [2,5]. Although not particularly covered in this article, the output from studies that provide the capacity to move beyond the evaluation of rare variation will be of great importance in these endeavors [109]. In the field of PD, genome-wide association studies have convincingly identified more than 90 more common variants that are implicated in PD risk, offering insights into the molecular-genetic architecture of the disease at a broader scope [109]. Implementation of complementary approaches for parallel investigations of rare and common variants may be very powerful for predicting more precisely the risk of developing PD and other movement disorders in the future. However, paradigm-shifting big-data studies are currently transforming different forms of movement disorders at different paces [2]. To equitably realize the benefits of big-data studies across indications, coordinated efforts are necessary. One example of underrepresentation in impactful big-data advancements is dystonia, which has long attracted much less support for research than other movement disorders such as PD or ataxia [2]. We must support synergies to provide research data of patients with all types of movement disorders to analytical approaches that enhance convergence of multi-layer disease information, computer science, and proof-of-concept translation. Critical factors for success will be the mainstreaming of combined genomics and multi-omics applications in daily care and the unification of independent data-rich initiatives, thus maximizing the potential for precision medicine. It is now a crucial time to demonstrate how big data can transform evidence-based diagnostics and therapy, enhancing decision-making and optimizing clinical impact at a global level.

CRediT authorship contribution statement

Alice Saparov: Writing – review & editing, Writing – original draft. **Michael Zech:** Writing – review & editing, Writing – original draft.

3. Disclosures

None of the authors report disclosures concerning the present manuscript.

Declaration of competing interest

The authors state that they have nothing to declare.

Acknowledgements

M.Z. is 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). M.Z.'s research is also supported by a "*Schlüsselprojekt*" grant from the Else Kröner-Fresenius-Stiftung (2022_EKSE.185). In addition, M.Z. receives 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. M.Z. has received research support from the German Research Foundation (DFG 458949627; ZE 1213/2-1). A.S. is supported by the Technical University of Munich - Institute for Advanced Study. The authors acknowledge the use of BioRender that was used to create parts of Figs. 1 and 2 (https://www.biorender.com/).

References

[1] L.F. Johansson, S. Laurie, D. Spalding, S. Gibson, D. Ruvolo, C. Thomas, D. Piscia, F. de Andrade, G. Been, M. Bijlsma, H. Brunner, S. Cimerman, F.Y. Dizjikan, K. Ellwanger, M. Fernandez, M. Freeberg, G.J. van de Geijn, R. Kanninga, V. Maddi, M. Mehtarizadeh, P. Neerincx, S. Ossowski, A. Rath, D. Roelofs-Prins, M. Stok-Benjamins, K.J. van der Velde, C. Veal, G. van der Vries, M. Wadsley, G. Warren, B. Zurek, T. Keane, H. Graessner, S. Beltran, M.A. Swertz, A. J. Brookes, R.D.c. Solve, An interconnected data infrastructure to support large-scale rare disease research, GigaScience 13 (2024).

[2] M. Zech, J. Winkelmann, Next-generation sequencing and bioinformatics in rare movement disorders, Nat. Rev. Neurol. 20 (2) (2024) 114–126.

[3] S.Y. Lim, A.H. Tan, A. Ahmad-Annuar, N.U. Okubadejo, K. Lohmann, H.R. Morris, T.S. Toh, Y.W. Tay, L.M. Lange, S. Bandres-Ciga, I. Mata, J.N. Foo, E. Sammler, J. C.E. Ooi, A.J. Noyce, N. Bahr, W. Luo, R. Ojha, A.B. Singleton, C. Blauwendraat, C. Klein, Uncovering the genetic basis of Parkinson's disease globally: from discoveries to the clinic, Lancet Neurol. (2024).

[4] E. Indelicato, S. Boesch, GAA/FGF14 ataxia: an ode to the phenotype-first approach, EBioMedicine 103 (2024) 105131.

[5] K.M. Boycott, L.P. Lau, C.M. Cutillo, C.P. Austin, International collaborative actions and transparency to understand, diagnose, and develop therapies for rare diseases, EMBO Mol. Med. 11 (5) (2019).

[6] S. Lunke, S.E. Bouffler, C.V. Patel, S.A. Sandaradura, M. Wilson, J. Pinner, M. F. Hunter, C.P. Barnett, M. Wallis, B. Kamien, T.Y. Tan, M.L. Freckmann, B. Chong, D. Phelan, D. Francis, K.S. Kassahn, T. Ha, S. Gao, P. Arts, M. R. Jackson, H.S. Scott, S. Eggers, S. Rowley, K. Boggs, A. Rakonjac, G.R. Brett, M. G. de Silva, A. Springer, M. Ward, K. Stallard, C. Simons, T. Conway, A. Halman, N.J. Van Bergen, T. Sikora, L.N. Semcesen, D.A. Stroud, A.G. Compton, D. R. Thorburn, K.M. Bell, S. Sadedin, K.N. North, J. Christodoulou, Z. Stark, Integrated multi-omics for rapid rare disease diagnosis on a national scale, Nat. Med. 29 (7) (2023) 1681–1691.

[7] K.M. Boycott, T. Hartley, K.D. Kernohan, D.A. Dyment, H. Howley, A.M. Innes, F. P. Bernier, M. Brudno, C. Care4Rare Canada, Care4Rare Canada: outcomes from a decade of network science for rare disease gene discovery, Am. J. Hum. Genet. 109 (11) (2022) 1947–1959.

[8] J.B. Byrd, A.C. Greene, D.V. Prasad, X. Jiang, C.S. Greene, Responsible, practical genomic data sharing that accelerates research, Nat. Rev. Genet. 21 (10) (2020) 615–629.

[9] D. Smirnov, N. Konstantinovskiy, H. Prokisch, Integrative omics approaches to advance rare disease diagnostics, J. Inherit. Metab. Dis. 46 (5) (2023) 824–838.

[10] K.M. Boycott, A. Rath, J.X. Chong, T. Hartley, F.S. Alkuraya, G. Baynam, A. J. Brookes, M. Brudno, A. Carracedo, J.T. den Dunnen, S.O.M. Dyke, X. Estivill, J. Goldblatt, C. Gonthier, S.C. Groft, I. Gut, A. Hamosh, P. Hieter, S. Hohn, M. E. Hurles, P. Kaufmann, B.M. Knoppers, J.P. Krischer, M. Macek Jr., G. Matthijs, A. Olry, S. Parker, J. Paschall, A.A. Philippakis, H.L. Rehm, P.N. Robinson, P. C. Sham, R. Stefanov, D. Taruscio, D. Unni, M.R. Vanstone, F. Zhang, H. Brunner, M.J. Bamshad, H. Lochmuller, International cooperation to enable the diagnosis of all rare genetic diseases, Am. J. Hum. Genet. 100 (5) (2017) 695–705.

[11] G. Eraslan, Z. Avsec, J. Gagneur, F.J. Theis, Deep learning: new computational modelling techniques for genomics, Nat. Rev. Genet. 20 (7) (2019) 389–403.

[12] K. Kerr, H. McAneney, L.J. Smyth, C. Bailie, S. McKee, A.J. McKnight, A scoping review and proposed workflow for multi-omic rare disease research, Orphanet J. Rare Dis. 15 (1) (2020) 107.

[13] B. Tesi, C. Boileau, K.M. Boycott, G. Canaud, M. Caulfield, D. Choukair, S. Hill, M. Spielmann, A. Wedell, V. Wirta, A. Nordgren, A. Lindstrand, Precision medicine in rare diseases: what is next? J. Intern. Med. 294 (4) (2023) 397–412.

[14] H.L. Rehm, Evolving health care through personal genomics, Nat. Rev. Genet. 18 (4) (2017) 259–267.

[15] M. Osmond, T. Hartley, D.A. Dyment, K.D. Kernohan, M. Brudno, O.J. Buske, A. M. Innes, K.M. Boycott, C. Care4Rare Canada, Outcome of over 1500 matches through the Matchmaker Exchange for rare disease gene discovery: the 2-year experience of Care4Rare Canada, Genet. Med. 24 (1) (2022) 100–108.

[16] Z. Stark, D. Glazer, O. Hofmann, A. Rendon, C.R. Marshall, G.S. Ginsburg, C. Lunt, N. Allen, M. Effingham, J. Hastings Ward, S.L. Hill, R. Ali, P. Goodhand, A. Page, H.L. Rehm, K.N. North, R.H. Scott, A call to action to scale up research and clinical genomic data sharing, Nat. Rev. Genet. (2024).

[17] K.J. Karczewski, M.P. Snyder, Integrative omics for health and disease, Nat. Rev. Genet. 19 (5) (2018) 299–310.

[18] S. Maiella, A. Olry, M. Hanauer, V. Lanneau, H. Lourghi, B. Donadille, C. Rodwell, S. Kohler, D. Seelow, S. Jupp, H. Parkinson, T. Groza, M. Brudno, P.N. Robinson, A. Rath, Harmonising phenomics information for a better interoperability in the rare disease field, Eur. J. Med. Genet. 61 (11) (2018) 706–714.

[19] K. Kochinke, C. Zweier, B. Nijhof, M. Fenckova, P. Cizek, F. Honti, S. Keerthikumar, M.A. Oortveld, T. Kleefstra, J.M. Kramer, C. Webber, M. A. Huynen, A. Schenck, Systematic phenomics analysis deconvolutes genes mutated in intellectual disability into biologically coherent modules, Am. J. Hum. Genet. 98 (1) (2016) 149–164.

[20] D.M. Fowler, H.L. Rehm, Will variants of uncertain significance still exist in 2030? Am. J. Hum. Genet. 111 (1) (2024) 5–10.

[21] M. Zech, R. Jech, S. Boesch, M. Skorvanek, S. Weber, M. Wagner, C. Zhao, A. Jochim, J. Necpal, Y. Dincer, K. Vill, F. Distelmaier, M. Stoklosa, M. Krenn, S. Grunwald, T. Bock-Bierbaum, A. Fecikova, P. Havrankova, J. Roth, I. Prihodova, M. Adamovicova, O. Ulmanova, K. Bechyne, P. Danhofer, B. Vesely, V. Han, P. Pavelekova, Z. Gdovinova, T. Mantel, T. Meindl, A. Sitzberger, S. Schroder, A. Blaschek, T. Roser, M.V. Bonfert, E. Haberlandt, B. Plecko, B. Leineweber, S. Berweck, T. Herberhold, B. Langguth, J. Svantnerova, M. Minar, G.A. Ramos-Rivera, M.H. Wojcik, S. Pajusalu, K. Ounap, U.A. Schatz, L. Polsler, I. Milenkovic, F. Laccone, V. Pilshofer, R. Colombo, S. Patzer, A. Iuso, J. Vera, M. Troncoso, F. Fang, H. Prokisch, F. Wilbert, M. Eckenweiler, E. Graf, D. S. Westphal, K.M. Riedhammer, T. Brunet, B. Alhaddad, R. Berutti, T.M. Strom, M. Hecht, M. Baumann, M. Wolf, A. Telegrafi, R.E. Person, F.M. Zamora, L. B. Henderson, D. Weise, T. Musacchio, J. Volkmann, A. Szuto, J. Becker, K. Cremer, T. Sycha, F. Zimprich, V. Kraus, C. Makowski, P. Gonzalez-Alegre, T. M. Bardakjian, L.J. Ozelius, A. Vetro, R. Guerrini, E. Maier, I. Borggraefe, A. Kuster, S.B. Wortmann, A. Hackenberg, R. Steinfeld, B. Assmann, C. Staufner, T. Opladen, E. Ruzicka, R.D. Cohn, D. Dyment, W.K. Chung, H. Engels, A. Ceballos-Baumann, R. Ploski, O. Daumke, B. Haslinger, V. Mall, K. Oexle, J. Winkelmann, Monogenic variants in dystonia: an exome-wide sequencing study, Lancet Neurol. 19 (11) (2020) 908–918.

[22] B.H. Laabs, K. Lohmann, E.J. Vollstedt, T. Reinberger, L.M. Nuxoll, G. Kilic-Berkmen, J.S. Perlmutter, S. Loens, C. Cruchaga, A. Franke, V. Dobricic, F. Hinrichs, A. Grozinger, E. Altenmuller, S. Bellows, S. Boesch, S.B. Bressman, K. R. Duque, A.J. Espay, A. Ferbert, J.S. Feuerstein, S. Frank, T. Gasser, B. Haslinger, R. Jech, F. Kaiser, C. Kamm, K. Kollewe, A.A. Kuhn, M.S. LeDoux, E. Lohmann, A. Mahajan, A. Munchau, T. Multhaupt-Buell, A. Pantelyat, S.E. Pirio Richardson, D. Raymond, S.G. Reich, R. Saunders Pullman, B. Schormair, N. Sharma, A. H. Sichani, K. Simonyan, J. Volkmann, A. Wagle Shukla, J. Winkelmann, L. J. Wright, M. Zech, K.E. Zeuner, S. Zittel, M. Kasten, Y.V. Sun, T. Baumer, N. Bruggemann, L.J. Ozelius, H.A. Jinnah, C. Klein, I.R. Konig, Genetic risk factors in isolated dystonia escape genome-wide association studies, Mov. Disord. 39 (11) (2024) 2110–2116.

[23] T. Klockgether, M. Synofzik, A.G.I.w.g.o. Coas, Registries, consensus recommendations for clinical outcome assessments and registry development in ataxias: ataxia global initiative (AGI) working group expert guidance, Cerebellum 23 (3) (2024) 924–930.

[24] A.B. Popejoy, S.M. Fullerton, Genomics is failing on diversity, Nature 538 (7624) (2016) 161–164.

[25] A.F. Schumacher-Schuh, A. Bieger, O. Okunoye, K.Y. Mok, S.Y. Lim, S. Bardien, A. Ahmad-Annuar, B.L. Santos-Lobato, M.Z. Strelow, M. Salama, S.C. Rao, Y. Z. Zewde, S. Dindayal, J. Azar, L.K. Prashanth, R. Rajan, A.J. Noyce, N. Okubadejo, M. Rizig, S. Lesage, I.F. Mata, P. Global Parkinson's Genetics, Underrepresented populations in Parkinson's genetics research: current landscape and future directions, Mov. Disord. 37 (8) (2022) 1593–1604.

[26] C.F. Wright, D.R. FitzPatrick, H.V. Firth, Paediatric genomics: diagnosing rare disease in children, Nat. Rev. Genet. 19 (5) (2018) 325.

[27] A. Hamosh, A.F. Scott, J.S. Amberger, C.A. Bocchini, V.A. McKusick, Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders, Nucleic Acids Res. 33 (Database issue) (2005) D514–D517.

[28] A.J.M. Dingemans, S. Jansen, J. van Reeuwijk, N. de Leeuw, R. Pfundt, J. Schuurs-Hoeijmakers, B.W. van Bon, C. Marcelis, C.W. Ockeloen, M. Willemsen, P.J. van der Sluijs, G.W.E. Santen, R.F. Kooy, A.T. Vulto-van Silfhout, T. Kleefstra, D.A. Koolen, L. Vissers, B.B.A. de Vries, Prevalence of comorbidities in individuals with neurodevelopmental disorders from the aggregated phenomics data of 51,227 pediatric individuals, Nat. Med. 30 (7) (2024) 1994–2003.

[29] S. Srivastava, S.A. Lewis, M.C. Kruer, A. Poduri, Underrepresentation of the term cerebral palsy in clinical genetics databases, Am. J. Med. Genet. 188 (12) (2022) 3555–3557.
[30] D. Rots, T.E. Jakub, C. Keung, A. Jackson, S. Banka, R. Pfundt, B.B.A. de Vries, R. H. van Jaarsveld, S.M.J. Hopman, E. van Binsbergen, I. Valenzuela, M. Hempel, T. Bierhals, F. Kortum, F. Lecoquierre, A. Goldenberg, J.M. Hertz, C.B. Andersen, M. Kibaek, E.J. Prijoles, R.E. Stevenson, D.B. Everman, W.G. Patterson, L. Meng, C. Gijavanekar, K. De Dios, S. Lakhani, T. Levy, M. Wagner, D. Wieczorek, P. J. Benke, M.S. Lopez Garcia, R. Perrier, S.B. Sousa, P.M. Almeida, M.J. Simoes, B. Isidor, W. Deb, A.A. Schmanski, O. Abdul-Rahman, C. Philippe, A.L. Bruel, L. Faivre, A. Vitobello, C. Thauvin, J.J. Smits, L. Garavelli, S.G. Caraffi, F. Peluso, L. Davis-Keppen, D. Platt, E. Royer, L. Leeuwen, M. Sinnema, A.P.A. Stegmann, C. Stumpel, G.E. Tiller, D.G.M. Bosch, S.T. Potgieter, S. Joss, M. Splitt, S. Holden, M. Prapa, N. Foulds, S. Douzgou, K. Puura, R. Waltes, A.G. Chiocchetti, C. M. Freitag, F.K. Satterstrom, S. De Rubeis, J. Buxbaum, B.D. Gelb, A. Branko, I. Kushima, J. Howe, S.W. Scherer, A. Arado, C. Baldo, O. Patat, D. Benedicte, D. Lopergolo, F.M. Santorelli, T.B. Haack, A. Dufke, M. Bertrand, R.J. Falb, A. Riess, P. Krieg, S. Spranger, M.F. Bedeschi, M. Iascone, S. Josephi-Taylor, T. Roscioli, M.F. Buckley, J. Liebelt, A.I. Dagli, E. Aten, A.C.E. Hurst, A. Hicks, M. Suri, E. Aliu, S. Naik, R. Sidlow, J. Coursimault, G. Nicolas, H. Kupper, F. Petit, V. Ibrahim, D. Top, F. Di Cara, C. Genomics England Research, R.J. Louie, E. Stolerman, H.G. Brunner, L. Vissers, J.M. Kramer, T. Kleefstra, The clinical and molecular spectrum of the KDM6B-related neurodevelopmental disorder, Am. J. Hum. Genet. 110 (6) (2023) 963–978.
[31] F. De Paoli, S. Berardelli, I. Limongelli, E. Rizzo, S. Zucca, VarChat: the generative AI assistant for the interpretation of human genomic variations, Bioinformatics 40 (4) (2024).
[32] N. Alnazzawi, P. Thompson, R. Batista-Navarro, S. Ananiadou, Using text mining techniques to extract phenotypic information from the PhenoCHF corpus, BMC Med. Inf. Decis. Making 15 (Suppl 2) (2015) S3 (Suppl 2).
[33] F.C.P. Navarro, H. Mohsen, C. Yan, S. Li, M. Gu, W. Meyerson, M. Gerstein, Genomics and data science: an application within an umbrella, Genome Biol. 20 (1) (2019) 109.
[34] H. Liang, B.Y. Tsui, H. Ni, C.C.S. Valentim, S.L. Baxter, G. Liu, W. Cai, D. S. Kermany, X. Sun, J. Chen, L. He, J. Zhu, P. Tian, H. Shao, L. Zheng, R. Hou, S. Hewett, G. Li, P. Liang, X. Zang, Z. Zhang, L. Pan, H. Cai, R. Ling, S. Li, Y. Cui, S. Tang, H. Ye, X. Huang, W. He, W. Liang, Q. Zhang, J. Jiang, W. Yu, J. Gao, W. Ou, Y. Deng, Q. Hou, B. Wang, C. Yao, Y. Liang, S. Zhang, Y. Duan, R. Zhang, S. Gibson, C.L. Zhang, O. Li, E.D. Zhang, G. Karin, N. Nguyen, X. Wu, C. Wen, J. Xu, W. Xu, B. Wang, W. Wang, J. Li, B. Pizzato, C. Bao, D. Xiang, W. He, S. He, Y. Zhou, W. Haw, M. Goldbaum, A. Tremoulet, C.N. Hsu, H. Carter, L. Zhu, K. Zhang, H. Xia, Evaluation and accurate diagnoses of pediatric diseases using artificial intelligence, Nat. Med. 25 (3) (2019) 433–438.
[35] M.M. Clark, A. Hildreth, S. Batalov, Y. Ding, S. Chowdhury, K. Watkins, K. Ellsworth, B. Camp, C.I. Kint, C. Yacoubian, L. Farnaes, M.N. Bainbridge, C. Beebe, J.J.A. Braun, M. Bray, J. Carroll, J.A. Cakici, S.A. Caylor, C. Clarke, M. P. Creed, J. Friedman, A. Frith, R. Gain, M. Gaughran, S. George, S. Gilmer, J. Gleeson, J. Gore, H. Grunenwald, R.L. Hovey, M.L. Janes, K. Lin, P. D. McDonagh, K. McBride, P. Mulrooney, S. Nahas, D. Oh, A. Oriol, L. Puckett, Z. Rady, M.G. Reese, J. Ryu, L. Salz, E. Sanford, L. Stewart, N. Sweeney, M. Tokita, L. Van Der Kraan, S. White, K. Wigby, B. Williams, T. Wong, M. S. Wright, C. Yamada, P. Schols, J. Reynders, K. Hall, D. Dimmock, N. Veeraraghavan, T. Defay, S.F. Kingsmore, Diagnosis of genetic diseases in seriously ill children by rapid whole-genome sequencing and automated phenotyping and interpretation, Sci. Transl. Med. 11 (489) (2019).
[36] J.O.B. Jacobsen, C. Kelly, V. Cipriani, G.E. Research Consortium, C.J. Mungall, J. Reese, D. Danis, P.N. Robinson, D. Smedley, Phenotype-driven approaches to enhance variant prioritization and diagnosis of rare disease, Hum. Mutat. 43 (8) (2022) 1071–1081.
[37] G.P.P. Investigators, D. Smedley, K.R. Smith, A. Martin, E.A. Thomas, E. M. McDonagh, V. Cipriani, J.M. Ellingford, G. Arno, A. Tucci, J. Vandrovcova, G. Chan, H.J. Williams, T. Ratnaike, W. Wei, K. Stirrups, K. Ibanez, L. Moutsianas, M. Wielscher, A. Need, M.R. Barnes, L. Vestito, J. Buchanan, S. Wordsworth, S. Ashford, K. Rehmstrom, E. Li, G. Fuller, P. Twiss, O. Spasic-Boskovic, S. Halsall, R.A. Floto, K. Poole, A. Wagner, S.G. Mehta, M. Gurnell, N. Burrows, R. James, C. Penkett, E. Dewhurst, S. Graf, R. Mapeta, M. Kasanicki, A. Haworth, H. Savage, M. Babcock, M.G. Reese, M. Bale, E. Baple, C. Boustred, H. Brittain, A. de Burca, M. Bleda, A. Devereau, D. Halai, E. Haraldsdottir, Z. Hyder, D. Kasperaviciute, C. Patch, D. Polychronopoulos, A. Matchan, R. Sultana, M. Ryten, A.L.T. Tavares, C. Tregidgo, C. Turnbull, M. Welland, S. Wood, C. Snow, E. Williams, S. Leigh, R. E. Foulger, L.C. Daugherty, O. Niblock, I.U.S. Leong, C.F. Wright, J. Davies, C. Crichton, J. Welch, K. Woods, L. Abulhoul, P. Aurora, D. Bockenhauer, A. Broomfield, M.A. Cleary, T. Lam, M. Dattani, E. Footitt, V. Ganesan, S. Grunewald, S. Compeyrot-Lacassagne, F. Muntoni, C. Pilkington, R. Quinlivan, N. Thapar, C. Wallis, L.R. Wedderburn, A. Worth, T. Bueser, C. Compton, C. Deshpande, H. Fassihi, E. Haque, L. Izatt, D. Josifova, S. Mohammed, L. Robert, S. Rose, D. Ruddy, R. Sarkany, G. Say, A.C. Shaw, A. Wolejko, B. Habib, G. Burns, S. Hunter, R.J. Grocock, S.J. Humphray, P.N. Robinson, M. Haendel, M. A. Simpson, S. Banka, J. Clayton-Smith, S. Douzgou, G. Hall, H.B. Thomas, R. T. O'Keefe, M. Michaelides, A.T. Moore, S. Malka, N. Pontikos, A.C. Browning, V. Straub, G.S. Gorman, R. Horvath, R. Quinton, A.M. Schaefer, P. Yu-Wai-Man, D.M. Turnbull, R. McFarland, R.W. Taylor, E. O'Connor, J. Yip, K. Newland, H. R. Morris, J. Polke, N.W. Wood, C. Campbell, C. Camps, K. Gibson, N. Koelling, T. Lester, A.H. Nemeth, C. Palles, S. Patel, N.B.A. Roy, A. Sen, J. Taylor, P. Cacheiro, J.O. Jacobsen, E.G. Seaby, V. Davison, L. Chitty, A. Douglas, K. Naresh, D. McMullan, S. Ellard, I.K. Temple, A.D. Mumford, G. Wilson, P. Beales, M. Bitner-Glindzicz, G. Black, J.R. Bradley, P. Brennan, J. Burn, P. F. Chinnery, P. Elliott, F. Flinter, H. Houlden, M. Irving, W. Newman, S. Rahman, J.A. Sayer, J.C. Taylor, A.R. Webster, A.O.M. Wilkie, W.H. Ouwehand, F. L. Raymond, J. Chisholm, S. Hill, D. Bentley, R.H. Scott, T. Fowler, A. Rendon, M. Caulfield 100, 000 genomes pilot on rare-disease diagnosis in health care - preliminary report, N. Engl. J. Med. 385 (20) (2021) 1868–1880.
[38] S. Kohler, N.C. Oien, O.J. Buske, T. Groza, J.O.B. Jacobsen, C. McNamara, N. Vasilevsky, L.C. Carmody, J.P. Gourdine, M. Gargano, J.A. McMurry, D. Danis, C.J. Mungall, D. Smedley, M. Haendel, P.N. Robinson, Encoding clinical data with the human phenotype Ontology for computational differential diagnostics, Curr Protoc Hum Genet 103 (1) (2019) e92.
[39] T. Brunet, R. Jech, M. Brugger, R. Kovacs, B. Alhaddad, G. Leszinski, K. M. Riedhammer, D.S. Westphal, I. Mahle, K. Mayerhanser, M. Skorvanek, S. Weber, E. Graf, R. Berutti, J. Necpal, P. Havrankova, P. Pavelekova, M. Hempel, U. Kotzaeridou, G.F. Hoffmann, S. Leiz, C. Makowski, T. Roser, S. A. Schroeder, R. Steinfeld, G. Strobl-Wildemann, J. Hoefele, I. Borggraefe, F. Distelmaier, T.M. Strom, J. Winkelmann, T. Meitinger, M. Zech, M. Wagner, De novo variants in neurodevelopmental disorders-experiences from a tertiary care center, Clin. Genet. 100 (1) (2021) 14–28.
[40] T.C. Hsieh, A. Bar-Haim, S. Moosa, N. Ehmke, K.W. Gripp, J.T. Pantel, M. Danyel, M.A. Mensah, D. Horn, S. Rosnev, N. Fleischer, G. Bonini, A. Hustinx, A. Schmid, A. Knaus, B. Javanmardi, H. Klinkhammer, H. Lesmann, S. Sivalingam, T. Kamphans, W. Meiswinkel, F. Ebstein, E. Kruger, S. Kury, S. Bezieau, A. Schmidt, S. Peters, H. Engels, E. Mangold, M. Kreiss, K. Cremer, C. Perne, R. C. Betz, T. Bender, K. Grundmann-Hauser, T.B. Haack, M. Wagner, T. Brunet, H. B. Bentzen, L. Averdunk, K.C. Coetzer, G.J. Lyon, M. Spielmann, C.P. Schaaf, S. Mundlos, M.M. Nothen, P.M. Krawitz, GestaltMatcher facilitates rare disease matching using facial phenotype descriptors, Nat. Genet. 54 (3) (2022) 349–357.
[41] A. Schmidt, M. Danyel, K. Grundmann, T. Brunet, H. Klinkhammer, T.C. Hsieh, H. Engels, S. Peters, A. Knaus, S. Moosa, L. Averdunk, F. Boschann, H.L. Sczakiel, S. Schwartzmann, M.A. Mensah, J.T. Pantel, M. Holtgrewe, A. Bosch, C. Weiss, N. Weinhold, A.A. Suter, C. Stoltenburg, J. Neugebauer, T. Kallinich, A.M. Kaindl, S. Holzhauer, C. Buhrer, P. Bufler, U. Kornak, C.E. Ott, M. Schulke, H.H. P. Nguyen, S. Hoffjan, C. Grasemann, T. Rothoeft, F. Brinkmann, N. Matar, S. Sivalingam, C. Perne, E. Mangold, M. Kreiss, K. Cremer, R.C. Betz, M. Mucke, L. Grigull, T. Klockgether, I. Spier, A. Heimbach, T. Bender, F. Brand, C. Stieber, A.M. Morawiec, P. Karakostas, V.S. Schafer, S. Bernsen, P. Weydt, S. Castro-Gomez, A. Aziz, M. Grobe-Einsler, O. Kimmich, X. Kobeleva, D. Onder, H. Lesmann, S. Kumar, P. Tacik, M.A. Basin, P. Incardona, M.A. Lee-Kirsch, R. Berner, C. Schuetz, J. Korholz, T. Kretschmer, N. Di Donato, E. Schrock, A. Heinen, U. Reuner, A.M. Hansske, F.J. Kaiser, E. Manka, M. Munteanu, A. Kuechler, K. Cordula, R. Hirtz, E. Schlapakow, C. Schlein, J. Lisfeld, C. Kubisch, T. Herget, M. Hempel, C. Weiler-Normann, K. Ullrich, C. Schramm, C. Rudolph, F. Rillig, M. Groffmann, A. Muntau, A. Tibelius, E.M.C. Schwaibold, C.P. Schaaf, M. Zawada, L. Kaufmann, K. Hinderhofer, P.M. Okun, U. Kotzaeridou, G. F. Hoffmann, D. Choukair, M. Bettendorf, M. Spielmann, A. Ripke, M. Pauly, A. Munchau, K. Lohmann, I. Huning, B. Hanker, T. Baumer, R. Herzog, Y. Hellenbroich, D.S. Westphal, T. Strom, R. Kovacs, K.M. Riedhammer, K. Mayerhanser, E. Graf, M. Brugger, J. Hoefele, K. Oexle, N. Mirza-Schreiber, R. Berutti, U. Schatz, M. Krenn, C. Makowski, H. Weigand, S. Schroder, M. Rohlfs, K. Vill, F. Hauck, I. Borggraefe, W. Muller-Felber, I. Kurth, M. Elbracht, C. Knopp, M. Begemann, F. Kraft, J.R. Lemke, J. Hentschel, K. Platzer, V. Strehlow, R. Abou Jamra, M. Kehrer, G. Demidov, S. Beck-Wodl, H. Graessner, M. Sturm, L. Zeltner, L.J. Schols, J. Magg, A. Bevot, C. Kehrer, N. Kaiser, E. Turro, D. Horn, A. Gruters-Kieslich, C. Klein, S. Mundlos, M. Nothen, O. Riess, T. Meitinger, H. Krude, P. M. Krawitz, T. Haack, N. Ehmke, M. Wagner, Next-generation phenotyping integrated in a national framework for patients with ultrarare disorders improves genetic diagnostics and yields new molecular findings, Nat. Genet. 56 (8) (2024) 1644–1653.
[42] J. Varghese, A. Brenner, M. Fujarski, C.M. van Alen, L. Plagwitz, T. Warnecke, Machine Learning in the Parkinson's disease smartwatch (PADS) dataset, NPJ Parkinsons Dis 10 (1) (2024) 9.
[43] J. Rexach, H. Lee, J.A. Martinez-Agosto, A.H. Nemeth, B.L. Fogel, Clinical application of next-generation sequencing to the practice of neurology, Lancet Neurol. 18 (5) (2019) 492–503.
[44] J.N. Foo, J.J. Liu, E.K. Tan, Whole-genome and whole-exome sequencing in neurological diseases, Nat. Rev. Neurol. 8 (9) (2012) 508–517.
[45] D.L. Fehlings, M. Zarrei, W. Engchuan, N. Sondheimer, B. Thiruvahindrapuram, J. R. MacDonald, E.J. Higginbotham, R. Thapa, T. Behlim, S. Aimola, L. Switzer, P. Ng, J. Wei, P.S. Danthi, G. Pellecchia, S. Lamoureux, K. Ho, S.L. Pereira, J. de Rijke, W.W.L. Sung, A. Mowjoodi, J.L. Howe, T. Nalpathamkalam, R. Manshaei, S. Ghaffari, J. Whitney, R.V. Patel, O. Hamdan, R. Shaath, B. Trost, S. Knights, D. Samdup, A. McCormick, C. Hunt, A. Kirton, A. Kawamura, R. Mesterman, J. W. Gorter, N. Dlamini, D. Merico, M. Hilali, K. Hirschfeld, K. Grover, N. X. Bautista, K. Han, C.R. Marshall, R.K.C. Yuen, P. Subbarao, M.B. Azad, S. E. Turvey, P. Mandhane, T.J. Moraes, E. Simons, G. Maxwell, M. Shevell, G. Costain, J.L. Michaud, F.F. Hamdan, J. Gauthier, K. Uguen, D.J. Stavropoulos, R.F. Wintle, M. Oskoui, S.W. Scherer, Comprehensive whole-genome sequence analyses provide insights into the genomic architecture of cerebral palsy, Nat. Genet. 56 (4) (2024) 585–594.
[46] Y. Wang, Y. Xu, C. Zhou, Y. Cheng, N. Qiao, Q. Shang, L. Xia, J. Song, C. Gao, Y. Qiao, X. Zhang, M. Li, C. Ma, Y. Fan, X. Peng, S. Wu, N. Lv, B. Li, Y. Sun, B. Zhang, T. Li, H. Li, J. Zhang, Y. Su, Q. Li, J. Yuan, L. Liu, A. Moreno-De-Luca, A. H. MacLennan, J. Gecz, D. Zhu, X. Wang, C. Zhu, Q. Xing, Exome sequencing reveals genetic heterogeneity and clinically actionable findings in children with cerebral palsy, Nat. Med. 30 (5) (2024) 1395–1405.

[47] M. Zech, R. Jech, S. Boesch, M. Skorvanek, J. Necpal, J. Svantnerova, M. Wagner, A. Sadr-Nabavi, F. Distelmaier, M. Krenn, T. Serranova, I. Rektorova, P. Havrankova, A. Mosejova, I. Prihodova, J. Sarlakova, K. Kulcsarova, O. Ulmanova, K. Bechyne, M. Ostrozovicova, V. Han, J.R. Ventosa, T. Brunet, R. Berutti, M. Shariati, A. Shoeibi, S.A. Schneider, A. Kuster, M. Baumann, D. Weise, F. Wilbert, W.G. Janzarik, M. Eckenweiler, V. Mall, B. Haslinger, S. Berweck, J. Winkelmann, K. Oexle, Scoring algorithm-based genomic testing in dystonia: a prospective validation study, Mov. Disord. 36 (8) (2021) 1959–1964.

[48] C.F. Wright, P. Campbell, R.Y. Eberhardt, S. Aitken, D. Perrett, S. Brent, P. Danecek, E.J. Gardner, V.K. Chundru, S.J. Lindsay, K. Andrews, J. Hampstead, J. Kaplanis, K.E. Samocha, A. Middleton, J. Foreman, R.J. Hobson, M.J. Parker, H. C. Martin, D.R. FitzPatrick, M.E. Hurles, H.V. Firth, D.D.D. Study, Genomic diagnosis of rare pediatric disease in the United Kingdom and Ireland, N. Engl. J. Med. 388 (17) (2023) 1559–1571.

[49] H. Stranneheim, K. Lagerstedt-Robinson, M. Magnusson, M. Kvarnung, D. Nilsson, N. Lesko, M. Engvall, B.M. Anderlid, H. Arnell, C.B. Johansson, M. Barbaro, E. Bjorck, H. Bruhn, J. Eisfeldt, C. Freyer, G. Grigelioniene, P. Gustavsson, A. Hammarsjo, M. Hellstrom-Pigg, E. Iwarsson, A. Jemt, M. Laaksonen, S. L. Enoksson, H. Malmgren, K. Naess, M. Nordenskjold, M. Oscarson, M. Pettersson, C. Rasi, A. Rosenbaum, E. Sahlin, E. Sardh, T. Stodberg, B. Tesi, E. Tham, H. Thonberg, V. Tohonen, U. von Dobeln, D. Vassiliou, S. Vonlanthen, A. C. Wikstrom, J. Wincent, O. Winqvist, A. Wredenberg, S. Ygberg, R. H. Zetterstrom, P. Marits, M.J. Soller, A. Nordgren, V. Wirta, A. Lindstrand, A. Wedell, Integration of whole genome sequencing into a healthcare setting: high diagnostic rates across multiple clinical entities in 3219 rare disease patients, Genome Med. 13 (1) (2021) 40.

[50] O. Abdelwahab, F. Belzile, D. Torkamaneh, Performance analysis of conventional and AI-based variant callers using short and long reads, BMC Bioinf. 24 (1) (2023) 472.

[51] M.A. DePristo, E. Banks, R. Poplin, K.V. Garimella, J.R. Maguire, C. Hartl, A. A. Philippakis, G. del Angel, M.A. Rivas, M. Hanna, A. McKenna, T.J. Fennell, A. M. Kernytsky, A.Y. Sivachenko, K. Cibulskis, S.B. Gabriel, D. Altshuler, M.J. Daly, A framework for variation discovery and genotyping using next-generation DNA sequencing data, Nat. Genet. 43 (5) (2011) 491–498.

[52] R. Poplin, P.C. Chang, D. Alexander, S. Schwartz, T. Colthurst, A. Ku, D. Newburger, J. Dijamco, N. Nguyen, P.T. Afshar, S.S. Gross, L. Dorfman, C. Y. McLean, M.A. DePristo, A universal SNP and small-indel variant caller using deep neural networks, Nat. Biotechnol. 36 (10) (2018) 983–987.

[53] M. Coutelier, M. Holtgrewe, M. Jager, R. Flottman, M.A. Mensah, M. Spielmann, P. Krawitz, D. Horn, D. Beule, S. Mundlos, Combining callers improves the detection of copy number variants from whole-genome sequencing, Eur. J. Hum. Genet. 30 (2) (2022) 178–186.

[54] S.R. Chintalaphani, S.S. Pineda, I.W. Deveson, K.R. Kumar, An update on the neurological short tandem repeat expansion disorders and the emergence of long-read sequencing diagnostics, Acta Neuropathol Commun 9 (1) (2021) 98.

[55] E. Dolzhenko, V. Deshpande, F. Schlesinger, P. Krusche, R. Petrovski, S. Chen, D. Emig-Agius, A. Gross, G. Narzisi, B. Bowman, K. Scheffler, J. van Vugt, C. French, A. Sanchis-Juan, K. Ibanez, A. Tucci, B.R. Lajoie, J.H. Veldink, F. L. Raymond, R.J. Taft, D.R. Bentley, M.A. Eberle, ExpansionHunter: a sequence-graph-based tool to analyze variation in short tandem repeat regions, Bioinformatics 35 (22) (2019) 4754–4756.

[56] H. Rafehi, J. Read, D.J. Szmulewicz, K.C. Davies, P. Snell, L.G. Fearnley, L. Scott, M. Thomsen, G. Gillies, K. Pope, M.F. Bennett, J.E. Munro, K.J. Ngo, L. Chen, M. J. Wallis, E.G. Butler, K.R. Kumar, K.H. Wu, S.E. Tomlinson, S. Tisch, A. Malhotra, M. Lee-Archer, E. Dolzhenko, M.A. Eberle, L.J. Roberts, B.L. Fogel, N. Bruggemann, K. Lohmann, M.B. Delatycki, M. Bahlo, P.J. Lockhart, An intronic GAA repeat expansion in FGF14 causes the autosomal-dominant adult-onset ataxia SCA50/ATX-FGF14, Am. J. Hum. Genet. 110 (1) (2023) 105–119.

[57] S.L. Amarasinghe, S. Su, X. Dong, L. Zappia, M.E. Ritchie, Q. Gouil, Opportunities and challenges in long-read sequencing data analysis, Genome Biol. 21 (1) (2020) 30.

[58] I. Stevanovski, S.R. Chintalaphani, H. Gamaarachchi, J.M. Ferguson, S.S. Pineda, C.K. Scriba, M. Tchan, V. Fung, K. Ng, A. Cortese, H. Houlden, C. Dobson-Stone, L. Fitzpatrick, G. Halliday, G. Ravenscroft, M.R. Davis, N.G. Laing, A. Fellner, M. Kennerson, K.R. Kumar, I.W. Deveson, Comprehensive genetic diagnosis of tandem repeat expansion disorders with programmable targeted nanopore sequencing, Sci. Adv. 8 (9) (2022) eabm5386.

[59] K. Daida, M. Funayama, K.J. Billingsley, L. Malik, A. Miano-Burkhardt, H. L. Leonard, M.B. Makarious, H. Iwaki, J. Ding, J.R. Gibbs, M. Ishiguro, H. Yoshino, K. Ogaki, G. Oyama, K. Nishioka, R. Nonaka, W. Akamatsu, C. Blauwendraat, N. Hattori, Long-read sequencing resolves a complex structural variant in PRKN Parkinson's disease, Mov. Disord. 38 (12) (2023) 2249–2257.

[60] K.J. Karczewski, L.C. Francioli, G. Tiao, B.B. Cummings, J. Alfoldi, Q. Wang, R. L. Collins, K.M. Laricchia, A. Ganna, D.P. Birnbaum, L.D. Gauthier, H. Brand, M. Solomonson, N.A. Watts, D. Rhodes, M. Singer-Berk, E.M. England, E.G. Seaby, J.A. Kosmicki, R.K. Walters, K. Tashman, Y. Farjoun, E. Banks, T. Poterba, A. Wang, C. Seed, N. Whiffin, J.X. Chong, K.E. Samocha, E. Pierce-Hoffman, Z. Zappala, A.H. O'Donnell-Luria, E.V. Minikel, B. Weisburd, M. Lek, J.S. Ware, C. Vittal, I.M. Armean, L. Bergelson, K. Cibulskis, K.M. Connolly, M. Covarrubias, S. Donnelly, S. Ferriera, S. Gabriel, J. Gentry, N. Gupta, T. Jeandet, D. Kaplan, C. Llanwarne, R. Munshi, S. Novod, N. Petrillo, D. Roazen, V. Ruano-Rubio, A. Saltzman, M. Schleicher, J. Soto, K. Tibbetts, C. Tolonen, G. Wade, M. E. Talkowski, C. Genome Aggregation Database, B.M. Neale, M.J. Daly, D. G. MacArthur, The mutational constraint spectrum quantified from variation in 141,456 humans, Nature 581 (7809) (2020) 434–443.

[61] E. Indelicato, A. Eberl, S. Boesch, L.M. Lange, C. Klein, K. Lohmann, M. Zech, Genome aggregation database version 4-allele frequency changes and impact on variant interpretation in dystonia, Mov. Disord. (2024).

[62] E. Poggio, L. Barazzuol, A. Salmaso, C. Milani, A. Deligiannopoulou, A.G. Cazorla, S.S. Jang, N. Julia-Palacios, B. Keren, R. Kopajtich, S.A. Lynch, C. Mignot, C. Moorwood, C. Neuhofer, V. Nigro, A. Oostra, H. Prokisch, V. Saillour, N. Schuermans, A. Torella, P. Verloo, E. Yazbeck, M. Zollino, R. Jech, J. Winkelmann, J. Necpal, T. Cali, M. Brini, M. Zech, ATP2B2 de novo variants as a cause of variable neurodevelopmental disorders that feature dystonia, ataxia, intellectual disability, behavioral symptoms, and seizures, Genet. Med. (2023) 100971.

[63] D.R. Tabet, D. Kuang, M.C. Lancaster, R. Li, K. Liu, J. Weile, A.G. Cote, Y. Wu, R. A. Hegele, D.M. Roden, F.P. Roth, Benchmarking computational variant effect predictors by their ability to infer human traits, Genome Biol. 25 (1) (2024) 172.

[64] J. Cheng, G. Novati, J. Pan, C. Bycroft, A. Zemgulyte, T. Applebaum, A. Pritzel, L. H. Wong, M. Zielinski, T. Sargeant, R.G. Schneider, A.W. Senior, J. Jumper, D. Hassabis, P. Kohli, Z. Avsec, Accurate proteome-wide missense variant effect prediction with AlphaMissense, Science 381 (6664) (2023) eadg7492.

[65] H. Gao, T. Hamp, J. Ede, J.G. Schraiber, J. McRae, M. Singer-Berk, Y. Yang, A.S. D. Dietrich, P.P. Fiziev, L.F.K. Kuderna, L. Sundaram, Y. Wu, A. Adhikari, Y. Field, C. Chen, S. Batzoglou, F. Aguet, G. Lemire, R. Reimers, D. Balick, M.C. Janiak, M. Kuhlwilm, J.D. Orkin, S. Manu, A. Valenzuela, J. Bergman, M. Rousselle, F. E. Silva, L. Agueda, J. Blanc, M. Gut, D. de Vries, I. Goodhead, R.A. Harris, M. Raveendran, A. Jensen, I.S. Chuma, J.E. Horvath, C. Hvilsom, D. Juan, P. Frandsen, F.R. de Melo, F. Bertuol, H. Byrne, I. Sampaio, I. Farias, J.V. do Amaral, M. Messias, M.N.F. da Silva, M. Trivedi, R. Rossi, T. Hrbek, N. Andriaholinirina, C.J. Rabarivola, A. Zaramody, C.J. Jolly, J. Phillips-Conroy, G. Wilkerson, C. Abee, J.H. Simmons, E. Fernandez-Duque, S. Kanthaswamy, F. Shiferaw, D. Wu, L. Zhou, Y. Shao, G. Zhang, J.D. Keyyu, S. Knauf, M.D. Le, E. Lizano, S. Merker, A. Navarro, T. Bataillon, T. Nadler, C.C. Khor, J. Lee, P. Tan, W.K. Lim, A.C. Kitchener, D. Zinner, I. Gut, A. Melin, K. Guschanski, M. H. Schierup, R.M.D. Beck, G. Umapathy, C. Roos, J.P. Boubli, M. Lek, S. Sunyaev, A. O'Donnell-Luria, H.L. Rehm, J. Xu, J. Rogers, T. Marques-Bonet, K.K. Farh, The landscape of tolerated genetic variation in humans and primates, Science 380 (6648) (2023) eabn8153.

[66] Y. Wei, T. Zhang, B. Wang, X. Jiang, F. Ling, M. Fang, X. Jin, Y. Bai, INDELpred: improving the prediction and interpretation of indel pathogenicity within the clinical genome, HGG Adv 5 (4) (2024) 100325.

[67] E. Giacopuzzi, N. Popitsch, J.C. Taylor, GREEN-DB: a framework for the annotation and prioritization of non-coding regulatory variants from whole-genome sequencing data, Nucleic Acids Res. 50 (5) (2022) 2522–2535.

[68] K. Jaganathan, S. Kyriazopoulou Panagiotopoulou, J.F. McRae, S.F. Darbandi, D. Knowles, Y.I. Li, J.A. Kosmicki, J. Arbelaez, W. Cui, G.B. Schwartz, E.D. Chow, E. Kanterakis, H. Gao, A. Kia, S. Batzoglou, S.J. Sanders, K.K. Farh, Predicting splicing from primary sequence with deep learning, Cell 176 (3) (2019) 535–548, e24.

[69] M.J. Landrum, J.M. Lee, M. Benson, G. Brown, C. Chao, S. Chitipiralla, B. Gu, J. Hart, D. Hoffman, J. Hoover, W. Jang, K. Katz, M. Ovetsky, G. Riley, A. Sethi, R. Tully, R. Villamarin-Salomon, W. Rubinstein, D.R. Maglott, ClinVar: public archive of interpretations of clinically relevant variants, Nucleic Acids Res. 44 (D1) (2016) D862–D868.

[70] D. Vitsios, R.S. Dhindsa, L. Middleton, A.B. Gussow, S. Petrovski, Prioritizing non-coding regions based on human genomic constraint and sequence context with deep learning, Nat. Commun. 12 (1) (2021) 1504.

[71] L.M. Crowther, M. Poms, B. Plecko, Multiomics tools for the diagnosis and treatment of rare neurological disease, J. Inherit. Metab. Dis. 41 (3) (2018) 425–434.

[72] A. Yang, M. Troup, J.W.K. Ho, Scalability and validation of big data bioinformatics software, Comput. Struct. Biotechnol. J. 15 (2017) 379–386.

[73] M.A. Levy, H. McConkey, J. Kerkhof, M. Barat-Houari, S. Bargiacchi, E. Biamino, M.P. Bralo, G. Cappuccio, A. Ciolfi, A. Clarke, B.R. DuPont, M.W. Elting, L. Faivre, T. Fee, R.S. Fletcher, F. Cherik, A. Foroutan, M.J. Friez, C. Gervasini, S. Haghshenas, B.A. Hilton, Z. Jenkins, S. Kaur, S. Lewis, R.J. Louie, S. Maitz, D. Milani, A.T. Morgan, R. Oegema, E. Ostergaard, N.R. Pallares, M. Piccione, S. Pizzi, A.S. Plomp, C. Poulton, J. Reilly, R. Relator, R. Rius, S. Robertson, K. Rooney, J. Rousseau, G.W.E. Santen, F. Santos-Simarro, J. Schijns, G.M. Squeo, M. St John, C. Thauvin-Robinet, G. Traficante, P.J. van der Sluijs, S.A. Vergano, N. Vos, K.K. Walden, D. Azmanov, T. Balci, S. Banka, J. Gecz, P. Henneman, J. A. Lee, M. Mannens, T. Roscioli, V. Siu, D.J. Amor, G. Baynam, E.G. Bend, K. Boycott, N. Brunetti-Pierri, P.M. Campeau, J. Christodoulou, D. Dyment, N. Esber, J.A. Fahrner, M.D. Fleming, D. Genevieve, K.D. Kerrnohan, A. McNeill, L.A. Menke, G. Merla, P. Prontera, C. Rockman-Greenberg, C. Schwartz, S. A. Skinner, R.E. Stevenson, A. Vitobello, M. Tartaglia, M. Alders, M.L. Tedder, B. Sadikovic, Novel diagnostic DNA methylation episignatures expand and refine the epigenetic landscapes of Mendelian disorders, HGG Adv 3 (1) (2022) 100075.

[74] E. Aref-Eshghi, J. Kerkhof, V.P. Pedro, D.I.F. Groupe, M. Barat-Houari, N. Ruiz-Pallares, J.C. Andrau, D. Lacombe, J. Van-Gils, P. Fergelot, C. Dubourg, V. Cormier-Daire, S. Rondeau, F. Lecoquierre, P. Saugier-Veber, G. Nicolas, G. Lesca, N. Chatron, D. Sanlaville, A. Vitobello, L. Faivre, C. Thauvin-Robinet, F. Laumonnier, M. Raynaud, M. Alders, M. Mannens, P. Henneman, R. C. Hennekam, G. Velasco, C. Francastel, D. Ulveling, A. Ciolfi, S. Pizzi, M. Tartaglia, S. Heide, D. Heron, C. Mignot, B. Keren, S. Whalen, A. Afenjar, T. Bienvenu, P.M. Campeau, J. Rousseau, M.A. Levy, L. Brick, M. Kozenko, T. B. Balci, V.M. Siu, A. Stuart, M. Kadour, J. Masters, K. Takano, T. Kleefstra, N. de Leeuw, M. Field, M. Shaw, J. Gecz, P.J. Ainsworth, H. Lin, D.I. Rodenhiser, M. J. Friez, M. Tedder, J.A. Lee, B.R. DuPont, R.E. Stevenson, S.A. Skinner, C.

E. Schwartz, D. Genevieve, B. Sadikovic, Evaluation of DNA methylation episignatures for diagnosis and phenotype correlations in 42 mendelian neurodevelopmental disorders, Am. J. Hum. Genet. 106 (3) (2020) 356–370.
[75] N. Mirza-Schreiber, M. Zech, R. Wilson, T. Brunet, M. Wagner, R. Jech, S. Boesch, M. Skorvanek, J. Necpal, D. Weise, S. Weber, B. Mollenhauer, C. Trenkwalder, E. M. Maier, I. Borggraefe, K. Vill, A. Hackenberg, V. Pilshofer, U. Kotzaeridou, E.M. C. Schwaibold, J. Hoefele, M. Waldenberger, C. Gieger, A. Peters, T. Meitinger, B. Schormair, J. Winkelmann, K. Oexle, Blood DNA methylation provides an accurate biomarker of KMT2B-related dystonia and predicts onset, Brain (2021).
[76] K. Oexle, M. Zech, L.G. Stuhn, S. Siegert, T. Brunet, W.M. Schmidt, M. Wagner, A. Schmidt, H. Engels, E. Tilch, O. Monestier, A. Destree, B. Hanker, S. Boesch, R. Jech, R. Berutti, F. Kaiser, B. Haslinger, T.B. Haack, B. Garavaglia, P. Krawitz, J. Winkelmann, N. Mirza-Schreiber, Episignature analysis of moderate effects and mosaics, Eur. J. Hum. Genet. 31 (9) (2023) 1032–1039.
[77] H. Prokisch, M. Zech, Proteomic profiling in dystonia: the next frontier for pathophysiology research and biomarker exploration, Mov. Disord. (2024).
[78] H. Lee, A.Y. Huang, L.K. Wang, A.J. Yoon, G. Renteria, A. Eskin, R.H. Signer, N. Dorrani, S. Nieves-Rodriguez, J. Wan, E.D. Douine, J.D. Woods, E. C. Dell'Angelica, B.L. Fogel, M.G. Martin, M.J. Butte, N.H. Parker, R.T. Wang, P. B. Shieh, D.A. Wong, N. Gallant, K.E. Singh, Y.J. Tavyev Asher, J.S. Sinsheimer, D. Krakow, S.K. Loo, P. Allard, J.C. Papp, N. Undiagnosed Diseases, C.G. S. Palmer, J.A. Martinez-Agosto, S.F. Nelson, Diagnostic utility of transcriptome sequencing for rare Mendelian diseases, Genet. Med. 22 (3) (2020) 490–499.
[79] V.A. Yepez, M. Gusic, R. Kopajtich, C. Mertes, N.H. Smith, C.L. Alston, R. Ban, S. Beblo, R. Berutti, H. Blessing, E. Ciara, F. Distelmaier, P. Freisinger, J. Haberle, S.J. Hayflick, M. Hempel, Y.S. Itkis, Y. Kishita, T. Klopstock, T.D. Krylova, C. Lamperti, D. Lenz, C. Makowski, S. Mosegaard, M.F. Muller, G. Munoz-Pujol, A. Nadel, A. Ohtake, Y. Okazaki, E. Procopio, T. Schwarzmayr, J. Smet, C. Staufner, S.L. Stenton, T.M. Strom, C. Terrile, F. Tort, R. Van Coster, A. Vanlander, M. Wagner, M. Xu, F. Fang, D. Ghezzi, J.A. Mayr, D. Piekutowska-Abramczuk, A. Ribes, A. Rotig, R.W. Taylor, S.B. Wortmann, K. Murayama, T. Meitinger, J. Gagneur, H. Prokisch, Clinical implementation of RNA sequencing for Mendelian disease diagnostics, Genome Med. 14 (1) (2022) 38.
[80] D.R. Murdock, H. Dai, L.C. Burrage, J.A. Rosenfeld, S. Ketkar, M.F. Muller, V. A. Yepez, J. Gagneur, P. Liu, S. Chen, M. Jain, G. Zapata, C.A. Bacino, H.T. Chao, P. Moretti, W.J. Craigen, N.A. Hanchard, N. Undiagnosed Diseases, B. Lee, Transcriptome-directed analysis for Mendelian disease diagnosis overcomes limitations of conventional genomic testing, J. Clin. Investig. 131 (1) (2021).
[81] F. Brechtmann, C. Mertes, A. Matuseviciute, V.A. Yepez, Z. Avsec, M. Herzog, D. M. Bader, H. Prokisch, J. Gagneur, OUTRIDER: a statistical method for detecting aberrantly expressed genes in RNA sequencing data, Am. J. Hum. Genet. 103 (6) (2018) 907–917.
[82] C. Mertes, I.F. Scheller, V.A. Yepez, M.H. Celik, Y. Liang, L.S. Kremer, M. Gusic, H. Prokisch, J. Gagneur, Detection of aberrant splicing events in RNA-seq data using FRASER, Nat. Commun. 12 (1) (2021) 529.
[83] T. Terkelsen, N.S. Mikkelsen, E.N. Bak, J. Vad-Nielsen, J. Blechingberg, S. Weiss, S.O. Drue, H. Andersen, B.S. Andresen, R.O. Bak, U.B. Jensen, CRISPR activation to characterize splice-altering variants in easily accessible cells, Am. J. Hum. Genet. 111 (2) (2024) 309–322.
[84] E.G. de Lope, R.T.J. Loo, A. Rauschenberger, M. Ali, L. Pavelka, T.M. Marques, C. P.C. Gomes, R. Kruger, E. Glaab, N.-P. Consortium, Comprehensive blood metabolomics profiling of Parkinson's disease reveals coordinated alterations in xanthine metabolism, NPJ Parkinsons Dis 10 (1) (2024) 68.
[85] C.D.M. van Karnebeek, S.B. Wortmann, M. Tarailo-Graovac, M. Langeveld, C. R. Ferreira, J.M. van de Kamp, C.E. Hollak, W.W. Wasserman, H.R. Waterham, R. A. Wevers, T.B. Haack, R.J.A. Wanders, K.M. Boycott, The role of the clinician in the multi-omics era: are you ready? J. Inherit. Metab. Dis. 41 (3) (2018) 571–582.
[86] I. Dzinovic, S. Boesch, M. Skorvanek, J. Necpal, J. Svantnerova, P. Pavelekova, P. Havrankova, E. Tsoma, E. Indelicato, E. Runkel, V. Held, D. Weise, W. Janzarik, M. Eckenweiler, S. Berweck, V. Mall, B. Haslinger, R. Jech, J. Winkelmann, M. Zech, Genetic overlap between dystonia and other neurologic disorders: a study of 1,100 exomes, Parkinsonism Relat. Disorders 102 (2022) 1–6.
[87] Y. Malakar, J. Lacey, N.A. Twine, R. McCrea, D.C. Bauer, Balancing the safeguarding of privacy and data sharing: perceptions of genomic professionals on patient genomic data ownership in Australia, Eur. J. Hum. Genet. 32 (5) (2024) 506–512.
[88] P.D. Stenson, M. Mort, E.V. Ball, K. Evans, M. Hayden, S. Heywood, M. Hussain, A.D. Phillips, D.N. Cooper, 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 (6) (2017) 665–677.
[89] C.M. Lill, A. Mashychev, C. Hartmann, K. Lohmann, C. Marras, A.E. Lang, C. Klein, L. Bertram, Launching the movement disorders society genetic mutation database (MDSGene), Mov. Disord. 31 (5) (2016) 607–609.
[90] G. Trieschmann, C. Wilhelm, S. Berweck, M. Zech, De novo retinoic acid receptor beta (RARB) variant associated with microphthalmia and dystonia, Eur. J. Med. Genet. 66 (8) (2023) 104802.
[91] V. Pitz, M.B. Makarious, S. Bandres-Ciga, H. Iwaki, T. andMe Research, A. B. Singleton, M. Nalls, K. Heilbron, C. Blauwendraat, Analysis of rare Parkinson's disease variants in millions of people, NPJ Parkinsons Dis 10 (1) (2024) 11.
[92] A.A. Philippakis, D.R. Azzariti, S. Beltran, A.J. Brookes, C.A. Brownstein, M. Brudno, H.G. Brunner, O.J. Buske, K. Carey, C. Doll, S. Dumitriu, S.O. Dyke, J. T. den Dunnen, H.V. Firth, R.A. Gibbs, M. Girdea, M. Gonzalez, M.A. Haendel, A. Hamosh, I.A. Holm, L. Huang, M.E. Hurles, B. Hutton, J.B. Krier, A. Misyura, C. J. Mungall, J. Paschall, B. Paten, P.N. Robinson, F. Schiettecatte, N.L. Sobreira, G. J. Swaminathan, P.E. Taschner, S.F. Terry, N.L. Washington, S. Zuchner, K. M. Boycott, H.L. Rehm, The Matchmaker Exchange: a platform for rare disease gene discovery, Hum. Mutat. 36 (10) (2015) 915–921.
[93] N. Sobreira, F. Schiettecatte, D. Valle, A. Hamosh, GeneMatcher: a matching tool for connecting investigators with an interest in the same gene, Hum. Mutat. 36 (10) (2015) 928–930.
[94] P.R. Blackburn, F. Ebstein, T.C. Hsieh, M. Motta, F.C. Radio, J.C. Herkert, T. Rinne, I. Thiffault, M. Rapp, M. Alders, S. Maas, B. Gerard, T. Smol, C. Vincent-Delorme, B. Cogne, B. Isidor, M. Vincent, R. Bachmann-Gagescu, A. Rauch, P. Joset, G.B. Ferrero, A. Ciolfi, T. Husson, A.M. Guerrot, C. Bacino, C. Macmurdo, S.S. Thompson, J.A. Rosenfeld, L. Faivre, F.T. Mau-Them, W. Deb, V. Vignard, P. B. Agrawal, J.A. Madden, A. Goldenberg, F. Lecoquierre, M. Zech, H. Prokisch, J. Necpal, R. Jech, J. Winkelmann, M.T. Koprusakova, V. Konstantopoulou, J. R. Younce, M. Shinawi, C. Mighton, C. Fung, C.F. Morel, J. Lerner-Ellis, S. DiTroia, M. Barth, D. Bonneau, I. Krapels, A.P.A. Stegmann, V. van der Schoot, T. Brunet, C. Bussmann, C. Mignot, G. Zampino, S.B. Wortmann, J.A. Mayr, R. G. Feichtinger, T. Courtin, C. Ravelli, B. Keren, A. Ziegler, L. Hasadsri, P. N. Pichurin, E.W. Klee, K. Grand, P.A. Sanchez-Lara, E. Kruger, S. Bezieau, H. Klinkhammer, P.M. Krawitz, E.E. Eichler, M. Tartaglia, S. Kury, T. Wang, Loss-of-Function variants in CUL3 cause a syndromic neurodevelopmental disorder, Ann. Neurol. (2024).
[95] M.A. Freeberg, L.A. Fromont, T. D'Altri, A.F. Romero, J.I. Ciges, A. Jene, G. Kerry, M. Moldes, R. Ariosa, S. Bahena, D. Barrowdale, M.C. Barbero, D. Fernandez-Orth, C. Garcia-Linares, E. Garcia-Rios, F. Haziza, B. Juhasz, O.M. Llobet, G. Milla, A. Mohan, M. Rueda, A. Sankar, D. Shaju, A. Shimpi, B. Singh, C. Thomas, S. de la Torre, U. Uyan, C. Vasallo, P. Flicek, R. Guigo, A. Navarro, H. Parkinson, T. Keane, J. Rambla, The European genome-phenome archive in 2021, Nucleic Acids Res. 50 (D1) (2022) D980–D987.
[96] D. Beijer, B.L. Fogel, S. Beltran, M.C. Danzi, A.H. Nemeth, S. Zuchner, M. Synofzik, p.W.G. Agi Ataxia Ngs genomics, Standards of NGS data sharing and analysis in ataxias: recommendations by the NGS working group of the ataxia global initiative, Cerebellum (2023).
[97] P. Global, Parkinson's genetics, GP2: the global Parkinson's genetics program, Mov. Disord. 36 (4) (2021) 842–851.
[98] B. Zurek, K. Ellwanger, L. Vissers, R. Schule, M. Synofzik, A. Topf, R.M. de Voer, S. Laurie, L. Matalonga, C. Gilissen, S. Ossowski, P.A.C. t Hoen, A. Vitobello, J. M. Schulze-Hentrich, O. Riess, H.G. Brunner, A.J. Brookes, A. Rath, G. Bonne, G. Gumus, A. Verloes, N. Hoogerbrugge, T. Evangelista, T. Harmuth, M. Swertz, D. Spalding, A. Hoischen, S. Beltran, H. Graessner, R.D.c. Solve, Solve-RD: systematic pan-European data sharing and collaborative analysis to solve rare diseases, Eur. J. Hum. Genet. 29 (9) (2021) 1325–1331.
[99] K. Walker, D. Kalra, R. Lowdon, G. Chen, D. Molik, D.C. Soto, F. Dabbaghie, A. A. Khleifat, M. Mahmoud, L.F. Paulin, M.S. Raza, S.P. Pfeifer, D.P. Agustinho, E. Aliyev, P. Avdeyev, E.R. Barrozo, S. Behera, K. Billingsley, L.C. Chong, D. Choubey, W. De Coster, Y. Fu, A.R. Gener, T. Hefferon, D.M. Henke, W. Hops, A. Illarionova, M.D. Jochum, M. Jose, R.K. Kesharwani, S.R.R. Kolora, J. Kubica, P. Lakra, D. Lattimer, C.S. Liew, B.W. Lo, C. Lo, A. Lotter, S. Majidian, S. K. Mendem, R. Mondal, H. Ohmiya, N. Parvin, C. Peralta, C.L. Poon, R. Prabhakaran, M. Saitou, A. Sammi, P. Sanio, N. Sapoval, N. Syed, T. Treangen, G. Wang, T. Xu, J. Yang, S. Zhang, W. Zhou, F.J. Sedlazeck, B. Busby, The third international hackathon for applying insights into large-scale genomic composition to use cases in a wide range of organisms, F1000Res 11 (2022) 530.
[100] E. Tambuyzer, B. Vandendriessche, C.P. Austin, P.J. Brooks, K. Larsson, K.I. Miller Needleman, J. Valentine, K. Davies, S.C. Groft, R. Preti, T.I. Oprea, M. Prunotto, Therapies for rare diseases: therapeutic modalities, progress and challenges ahead, Nat. Rev. Drug Discov. 19 (2) (2020) 93–111.
[101] S.A. Lewis, M. Chopra, J.S. Cohen, J.M. Bain, B. Aravamuthan, J.B. Carmel, M. C. Fahey, R. Segel, R.F. Wintle, M. Zech, H. May, N. Haque, D. Fehlings, S. Srivastava, M.C. Kruer, Clinical actionability of genetic findings in cerebral palsy: a systematic review and meta-analysis, JAMA Pediatr. (2024).
[102] C.L. van Eyk, M.C. Fahey, J. Gecz, Redefining cerebral palsies as a diverse group of neurodevelopmental disorders with genetic aetiology, Nat. Rev. Neurol. 19 (9) (2023) 542–555.
[103] F.W. Pun, I.V. Ozerov, A. Zhavoronkov, AI-powered therapeutic target discovery, Trends Pharmacol. Sci. 44 (9) (2023) 561–572.
[104] A.V. Sadybekov, V. Katritch, Computational approaches streamlining drug discovery, Nature 616 (7958) (2023) 673–685.
[105] D. Ochoa, M. Karim, M. Ghoussaini, D.G. Hulcoop, E.M. McDonagh, I. Dunham, Human genetics evidence supports two-thirds of the 2021 FDA-approved drugs, Nat. Rev. Drug Discov. 21 (8) (2022) 551.
[106] C.S. Storm, D.A. Kia, M.M. Almramhi, S. Bandres-Ciga, C. Finan, C. International Parkinson's Disease Genomics, A.D. Hingorani, N.W. Wood, Finding genetically-supported drug targets for Parkinson's disease using Mendelian randomization of the druggable genome, Nat. Commun. 12 (1) (2021) 7342.
[107] S. Mittal, I. Tang, J.G. Gleeson, Evaluating human mutation databases for "treatability" using patient-customized therapy, Médica Sur 3 (11) (2022) 740–759.
[108] J. Kim, S. Woo, C.M. de Gusmao, B. Zhao, D.H. Chin, R.L. DiDonato, M.A. Nguyen, T. Nakayama, C.A. Hu, A. Soucy, A. Kuniholm, J.K. Thornton, O. Riccardi, D. A. Friedman, C.M. El Achkar, Z. Dash, L. Cornelissen, C. Donado, K.N.W. Faour, L. W. Bush, V. Suslovitch, C. Lentucci, P.J. Park, E.A. Lee, A. Patterson, A. A. Philippakis, B. Margus, C.B. Berde, T.W. Yu, A framework for individualized splice-switching oligonucleotide therapy, Nature 619 (7971) (2023) 828–836.
[109] C. Blauwendraat, M.A. Nalls, A.B. Singleton, The genetic architecture of Parkinson's disease, Lancet Neurol. 19 (2) (2020) 170–178.