Preclinical alternative drug discovery programs for monogenic rare diseases. Should small molecules or gene therapy be used? The case of hereditary spastic paraplegias

Matteo Rossi Sebastiano [1], Shinji Hadano [2], Fabrizia Cesca [3], IAHSP Consortium [†], Giuseppe Ermondi [1,*]

[1] University of Torino, Molecular Biotechnology and Health Sciences Department, CASSMedChem, Piazza Nizza, 10138 Torino, Italy
[2] Molecular Neuropathobiology Laboratory, Department of Physiology, Tokai University School of Medicine, Isehara, Japan
[3] Department of Life Sciences, University of Trieste, 34127 Trieste, Italy

Patients diagnosed with rare diseases and their and families search desperately to organize drug discovery campaigns. Alternative models that differ from default paradigms offer real opportunities. There are, however, no clear guidelines for the development of such models, which reduces success rates and raises costs. We address the main challenges in making the discovery of new preclinical treatments more accessible, using rare hereditary paraplegia as a paradigmatic case. First, we discuss the necessary expertise, and the patients' clinical and genetic data. Then, we revisit gene therapy, *de novo* drug development, and drug repurposing, discussing their applicability. Moreover, we explore a pool of recommended *in silico* tools for pathogenic variant and protein structure prediction, virtual screening, and experimental validation methods, discussing their strengths and weaknesses. Finally, we focus on successful case applications.

Keywords: Rare diseases; Hereditary spastic paraplegia; Drug repurposing; Gene therapy; Virtual screening; In silico drug discovery

Introduction: the changing landscape of drug discovery for rare diseases

Rare diseases impact at least 300–400 million individuals globally, often leading to chronic illness, disability, and premature mortality.[(p1)] Correct diagnoses are not always known[(p2),(p3)] and can be insufficient, as currently there are no treatments for 95% of the 7,000 identified rare diseases. Therefore, urgent solutions are necessary to enhance the possibility of discovering therapies for rare conditions.

It is widely recognized that conventional drug discovery pipelines face significant limitations, particularly in terms of timeliness and costs,[(p4)] that greatly hinder the development of treatments for rare diseases. The high costs associated with research and development for rare diseases are seldom supported by the pharmaceutical sector due to the limited potential for revenue generation.[(p5)] As a result, big pharma companies have shown little interest in rare diseases, necessitating specific and often unique pipelines for these diseases. Relatives, patient associations, and other collaborative networks who support research using alternative R&D models have found options to address these challenges.[(p6)] These alternative models have been developing organically for decades, although clear paradigms have not yet been established. It is worth noting that discussions with key stakeholders have revealed a lack of awareness regarding

* *Corresponding author. Ermondi, G. (giuseppe.ermondi@unito.it).*
† A list of the members and their affiliations is presented at the end of the paper.

1359-6446/© 2024 The Author(s). Published by Elsevier Ltd.This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
https://doi.org/10.1016/j.drudis.2024.104138

the various drug discovery strategies available to identify new pharmacological treatments. In addition, although there are shared steps among these approaches, there are also distinct steps that have not yet been formally documented.

We use Hereditary spastic paraplegia (HSP, ORPHAnet code 685) as an example of a rare monogenic disorder. HSPs are associated with mutations in multiple genetic loci, resulting in a wide range of clinical manifestations.[p7] As for most genetic disorders, new mutations that cause HSPs are continuously being discovered.[p8] Despite this complexity, most of the genes that are involved in these conditions encode proteins that converge into a relatively small group of pathways, presenting opportunities to understand common mechanisms and to develop treatments.[p9]

We aim to offer guidelines for establishing a preclinical drug discovery program for monogenic rare genetic HSPs, utilizing an alternative approach to that used by the pharmaceutical industry. We start by introducing the essential components of the program, namely, the actors with the required expertise and the patients' clinical and genetic data. We then outline various strategies for discovering new treatments and offer insights into the challenges that must be overcome in the process. Like previous reports,[p10] our paper places a particular emphasis and provides an update on the *in-silico* tools currently used in repurposing strategies, which are used in combination with traditional docking methods (all related links are reported in the 'Supporting information'). Next, we delve into the essential experimental requirements for validating computational hypotheses and highlight the most common issues associated with any drug candidate. Towards the conclusion, we provide two case studies to exemplify the application of these emerging paradigms. The first case study involves Infantile ascending hereditary spastic paralysis (IAHSP) and the recent discovery of a small molecule, menatetrenone, that has been approved for compassionate use in one patient with a specific missense mutation in the *ALS2* gene.[p11] The second case study revolves around a gene therapy treatment called MELPIDA, which is currently undergoing Phase 2 clinical trials for Hereditary spastic paraplegia 50 (SPG50).

The required team

A team dedicated to finding treatments for rare diseases should be multi- and interdisciplinary (Table 1). Patients' families and clinicians play a crucial role as they are often the first to record the symptoms exhibited by a patient. A diagnosis is typically formulated on the basis of these clinical observations and genetic screening. In many cases, however, the gene–disease association is still unknown and the role of a medical geneticist becomes vital in formulating a diagnosis. Experts in biology identify underlying pathogenic cellular mechanisms, whereas chemical and structural biologists are able to rationalize the associated molecular mechanisms. Their contributions provide insights into potential targets for therapeutic intervention. Having obtained vital molecular information, team members with expertise in medicinal chemistry are essential to design therapeutic strategies. Drug candidates are then tested in cell systems to assess their potential as the subjects of clinical testing. Finally, the large amount of data, deriving from online databases, computational predictions, clinical information, and experiments, often requires that the team includes a data scientist who ensures that all of the information is handled in a suitable way. Notably, experts from various fields contribute across different stages of the drug discovery process and individuals may perform multiple professional roles with, for example, the team leader being one of the experts.

This alternative drug discovery model faces several challenges, including the extensive specialization that often hampers effective communication within the team, mainly resulting from a limited understanding of each other's challenges and often exacerbated by geographical, language, and cultural barriers. To overcome these challenges, the team leader should maintain close contact with the patients' families and associations, who in turn support the research through communication and fundraising. Another challenge arises from the potentially conflicting priorities of academics and industry professionals. The formers typically focus on publishing their work, while the latter prioritize practical success. In our opinion, the publication of key findings should be treated as validation of the work necessary to achieve a therapy, rather than the ultimate goal of the program.

Gathering the necessary data

The process of therapy development typically begins with clinical data collection. Genetic analyses are conducted to identify gene variants, followed by assessments of pathogenicity. Newborn whole exome screening (WES) represents the most powerful tool to identify novel variants, but not all hospitals are equipped for it. Alternatively, screening tests for specific mutations represent the most widely available resource, although variants of rare genes that are not included in the standard panels are likely to escape detection. Much information to place a patient within

TABLE 1

Team members, required expertise, and roles in rare disease drug discovery.

Professional figure	Expertise	Role
Team leader	Scientific background	Coordination
Family and associations	Communication	Coordination and fundraising
Clinician	Medicine	Diagnosis, compassionate use proposal, administration, follow-up
Geneticist	Medical genetics	Genetic investigation and interpretation
Data scientist	IT/data sciences	Harmonization of datasets, pathogenicity prediction, and experimental data
Biologist	Cell Biology, drug testing	Elucidation cellular mechanisms; drug testing in cell systems
Chemical biologist	Biophysics	Investigation of pathogenic mechanisms at the molecular level
Structural biologist	Structural biology	Prediction of protein variant structures and experimental resolution
Medicinal chemist	Drug design, drug testing	Therapeutic strategy design and testing

the landscape of a specific rare disease can be retrieved from public sources.

Clinical data sources

The lack of epidemiological and clinical data on rare diseases poses challenges for health service planning and clinical research. Patient registries offer a potential solution[p12] by using observational studies and by collecting mostly longitudinal real-world data about a population with a specific disease in a specific geographic area (regional, national, and international).[p13] The final goal is to enhance the quality of patient care by improving health policies and conducting targeted clinical trials.[p14] Patient registries are also valuable tools for clinical research,[p15] as they provide an understanding of the natural history of diseases, enable the planning of clinical trials, and facilitate patient enrolment and the attainment of a sample that is large enough to allow further research.

There are currently several registries related to HSPs. The Treat-HSP registry (Table S1) collects clinical and biological data on HSPs. Its aims are to develop and validate outcome metrics for clinical trials and to identify shared pathways and novel therapeutic targets. Access to the registry's web platform requires personal credentials. The Registry and Natural History Study for Early Onset HSP focuses on gathering longitudinal clinical data and biological samples from patients who exhibited the onset of HSP symptoms at 18 years old or younger. Its goal is to advance understanding of the causes, clinical progression, diagnosis, and treatment of these conditions. Access to the data will be limited to the staff responsible for the recruitment and maintenance of the registry. In addition, the soon to be launched Italian national registry for HSPs, known as STOP-HSP.net, is inspired by the same principles (Table S1).

Gene variants that are associated with specific diseases

A necessary step in a drug discovery program is the gathering of information about known pathogenic variants. Several freely accessible online databases are available to researchers, and these typically include information such as type of variants, position on the reference genome, and classification according to the American College of Medical Genetics (ACMG) guidelines.[p16] In some cases, the database also includes a description of the patient's phenotype. Examples include ClinVar,[p17] which archives and consolidates information about the relationship between genetic variation and human health; the Leiden Open Variation Database (LOVD),[p18] a flexible tool for collecting genomic variants and phenotypes from both patient- and gene-centric perspectives; and the Human Gene Mutation Database (HGMD), which compiles all known published pathogenic gene variants that are responsible for inherited human diseases.[p19] These databases serve as initial resources for understanding the impact of a variant (Table S1).

Predicting variant pathogenicity

When a new gene variant is identified, it is crucial to determine whether it is disease causative. The first interpretation step is assessing the variant frequency. If commonly found in unaffected individuals, it is less likely to be disease-causing. One of the most used databases is the Genome Aggregation Database (gnomAD),[p20] which has been designed as comprehensive resource. The latest version, gnomAD 4.0,[p20] is one of the largest and most ethnically diverse population databases and includes data from over 800,000 individuals and millions of genetic variants.

The second step is to evaluate the variant and gene in the context of the patient's family history and clinical manifestations. Accurate and comprehensive clinical information is essential for interpreting novel gene variants. By testing other family members, particularly parents and siblings, it is possible to determine whether the variant co-segregates with the disease within the family. If the variant, or combination of variants depending on the Mendelian inheritance pattern, is found in affected family members but not in unaffected individuals, there is strong evidence of disease causality. In autosomal dominant diseases, *de novo* variants are more likely to be disease-causing. Therefore, it is recommended that 'trio testing', in which the parents are analyzed alongside the proband (i.e. the first individual in a family to be identified as possibly having a genetic condition), is carried out as part of gene-panel, exome, or genome sequencing.

The ACMG has developed comprehensive guidelines for the classification of sequence variants based on a combination of expert opinions and empirical data. They provide a systematic approach to interpreting genetic variants and assigning them to specific categories based on their potential impact on diseases.[p16],[p21] The classification system in the ACGM guidelines places variants within one of five tiers:[p16] pathogenic, likely pathogenic, uncertain significance, likely benign, and benign.[p16] The factors considered are the type of variant (e.g., missense, nonsense, splice site), functional studies, segregation within families, prevalence in affected individuals and the general population, and information from disease and population databases such as ClinVar, OMIM, DECIPHER, dbSNP, dbVar, and gnomAD (Table S1).

Furthermore, there are computational predictive tools, both publicly available and commercial, to interpret sequence variants. Such programs can predict the impact of sequence changes on nucleotides and amino acids, as well as on primary and alternative gene transcripts, on other genomic elements, and on the encoded protein itself. These tools can be divided into two main categories: those that determine whether a missense change is damaging or not (e.g., PolyPhen-2,[p22] SIFT,[p23] REVEL, and CADD[p24]) and those that predict whether there is an effect on splicing (e.g., GeneSplicer, NNSplice, SpliceAI, and CADD splicing).[p25],[p26] Last, internet-based tools, such as VarSome and Franklin by Genoox, allow more rapid calculation of pathogenicity scores on the basis of ACGM criteria.[p27],[p28] In general, it is recommended that multiple *in silico* prediction software tools are used together for variant interpretation, as the different tools have different strengths and weaknesses depending on the algorithms employed.

Three-dimensional structure of the protein

Missense changes in a protein encoded by a disease-associated gene can also be analyzed by studying the three-dimensional (3D) structure of the protein, which may reveal whether the mutation disrupts or modifies the protein's function. The availability of the 3D structure of the protein is crucial in this process.

The current version of UniProt[p29] includes a "Structure" menu that provides direct links to both experimental (from the Protein Data Bank) and computed (from the AlphaFold[p30] database) 3D protein structures. The significance of structural data and how these data can enhance the prediction of mutation effects will be discussed in detail in later sections.

Manual curation of patient datasets

All the above data, combined with information from the literature and other patients' descriptions, generate complexity. The resulting collection of data is often overwhelming, incomplete, and confusing. To better organize and standardize the data, it is beneficial to arrange the data in a structured framework, such as a dataset, that offers improved navigation of large data volumes. Some clinical studies have made efforts in this direction, but there are currently no unified organization criteria that encompass genetic, structural, and clinical features.[p31],[p32]

An ideal structure for mutation datasets (Figure 1) should contain: (a) an internal anonymized patient ID with relevant information such as country of origin and reference information (publication or clinical reference details); (b) detailed clinical information, containing a patient's clinical presentation, symptoms, medical history, and prediction from knowledge-based pathogenicity scores (e.g., CADD); (c) harmonized genetic information, including the specific gene(s) associated with the disease and the identified mutations with coherent reference genome and transcript references (retrieved from ClinVar, or Mutation Taster[p33]) (d) chemical and functional information such predicted Nonsense Mediated Decay and predictions of the effect of each variant on the 3D protein structure.

Strategies to identify a targeted pharmacological treatment

After collecting clinical, genetic and protein structural data, it is necessary to determine the most suitable drug discovery strategy for the studied disease. Some authors have proposed approaches based on Clinical Outcome Pathways, in which drugs are identified on the basis of their known effects on cellular pathways. As this approach has been described elsewhere,[p34] we focus here on targeted approaches. There are three main targeted approaches to address rare diseases: gene therapy is based on large molecules, whereas *de novo* drug development and drug repurposing are both based on small molecules (Figure 2).[p4]

Gene therapy, or gene replacement therapy, involves the delivery of normal copies of the disease-causing gene to the patient to restore the altered phenotype. This strategy is primarily used in the treatment of monogenic disorders.[p35] Adeno-associated viral vectors (AAVs) are commonly used as delivery vectors because of their favorable characteristics; they are relatively non-pathogenic and non-replicating, they are able to transduce both dividing and non-dividing cells, and they do not integrate into the host genome.[p35] Nevertheless, the limited availability of long-term evidence regarding the safety and efficacy of gene therapy is a concern. Moreover, there are ongoing debates about the ethical implications of such treatments. Gene therapy is expensive and often inaccessible to many patients who require it. Last, the size of the gene can affect the feasibility and affordability of this approach, with smaller genes being more suitable.[p36]

The development of new drugs for rare diseases, known as orphan drugs, involves the design of new small molecules (Figure 2). Despite the Orphan Drug Act of 1983, which has encouraged efforts in this direction, significant financial and time investments are still missing, mainly due to limited profitability. Drug repurposing, as well as drug repositioning or reprofiling, offers a more cost-effective and quicker alternative (Figure 2). The rationale behind this approach is that a single drug can have multiple targets and can exert multiple effects. Therefore, a drug that was initially approved to treat one condition may have beneficial effects in treating different disorders,[p37] and having already undergone safety testing in Phase I clinical trials, the risk of failure is relatively small when compared to that of a newly developed compound.[p38] Another opportunity involves repro-

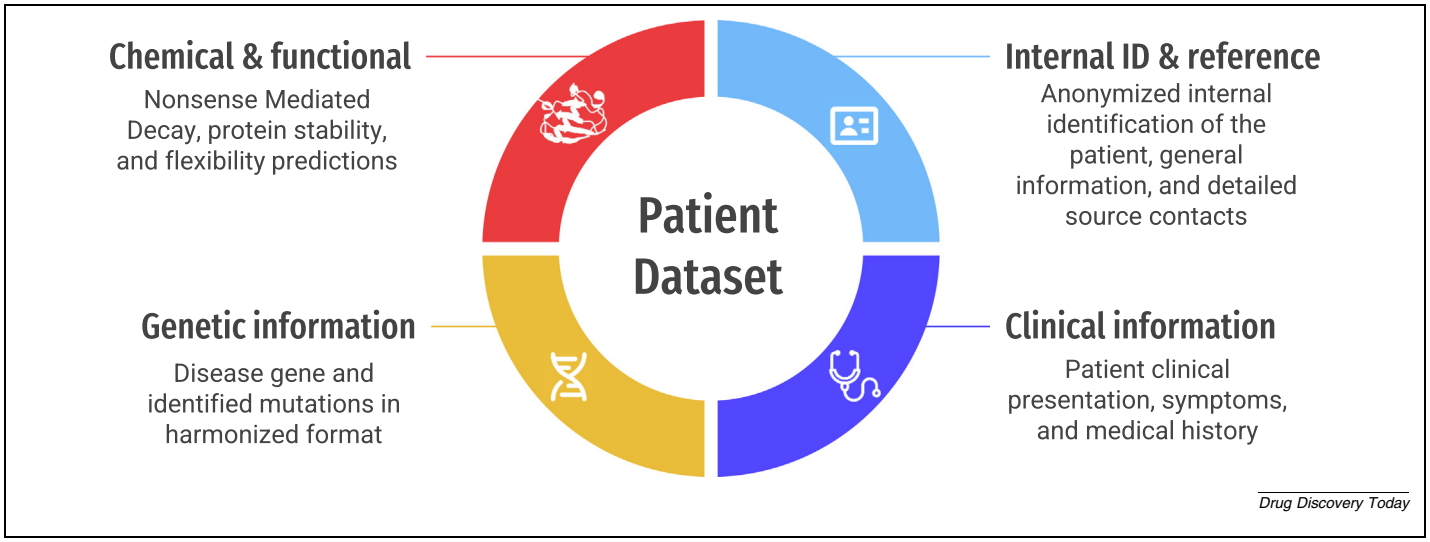

FIGURE 1

Four main categories of information that should be included in a patient dataset.

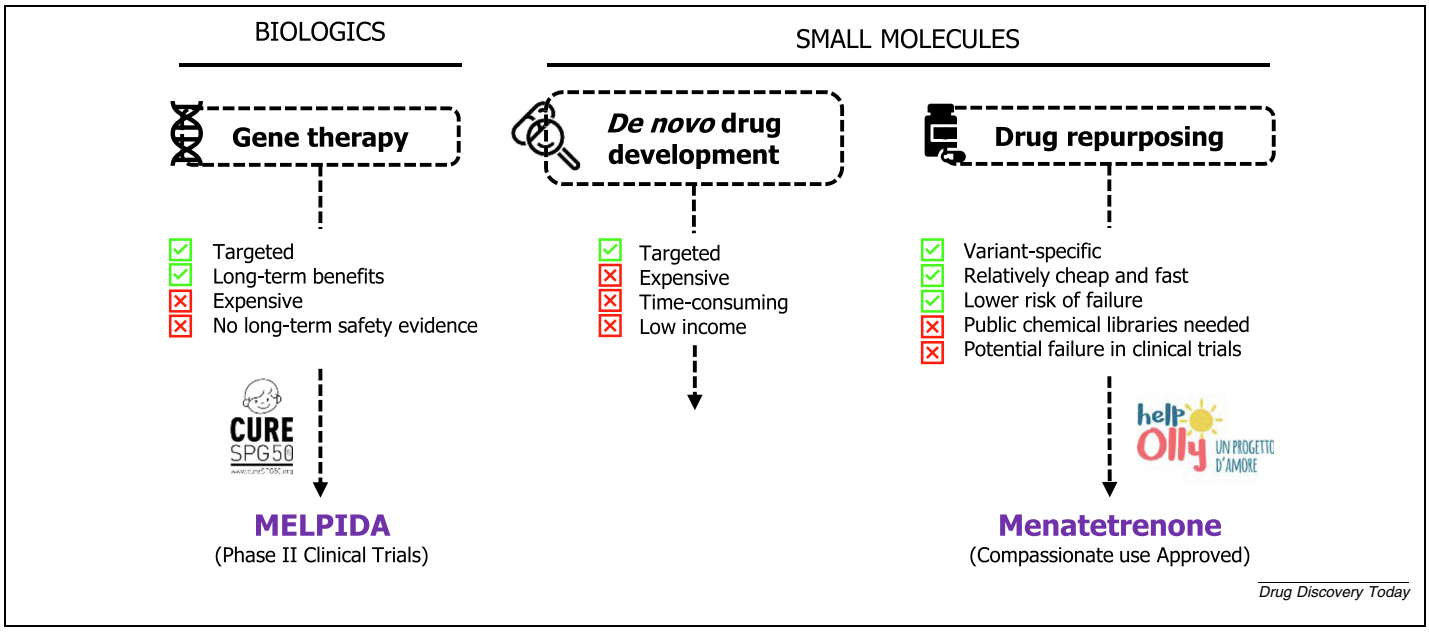

FIGURE 2

The three main development pipelines for pharmacological treatments of rare diseases: gene therapy, *de novo* development of a small molecule, and repurposing of a small molecule. Each approach has its own advantages and disadvantages. CureSPG50 and HelpOlly are two non-profit associations that have supported the development of gene therapy and repurposing approaches in the field of hereditary spastic paralysis (HSP).

filing drug candidates in clinical trials that have already passed safety testing. In this scenario, however, the higher risk of failure and the costs for Phase II, Phase III, or N-of-one clinical trials remains.[p39],[p40] Thus, the risk–benefit balance is an important aspect when considering a drug-repositioning strategy. Some molecules, including some classical chemotherapeutics, may pose a significant risk to the well-being of patients due to their side effects without providing commensurate benefits for rare disease patients. Repurposed drugs are often specific for individual patients as they are expected to treat a specific mutation, so each must be evaluated on a case-by-case basis.

Focus on repurposing: defining a strategy

Many authors have suggested that drug repurposing is a viable strategy in the field of rare diseases.[p4] Efforts have been made to list, collect, and organize resources that are useful when adopting this approach.[p41] Nevertheless, there is still a lack of a concise and definitive strategy for drug repurposing. The potential of computational approaches is often underestimated and limited to virtual screening, and protocols to assess the druggability of mutation sites are not commonly employed. Figure 3 schematizes the main steps of the computational pipeline, which we believe should precede experimental efforts in drug repurposing. The steps are: (a) look for altered protein products; (b) look at the evolutionary conservation of the changed amino acid; (c) determine the structural features of the mutation site; (d) assess the impact of the mutation on protein–protein interaction networks; (e) look for the presence of a potential binding pocket at the mutation site. Methods for performing these steps are found in Table 2 and discussed below, and a comprehensive list of tools is provided in Table S1.

The computational pipeline for drug repurposing starts with determining the stability of the aberrant mRNA, that is, by determining whether protein translation occurs. To assess this, predictors of Nonsense-mediated decay (NMD), a cellular quality-control mechanism that prevents the synthesis of harmful proteins, are used.[p42] Examples of such predictors include NMDesc[p43] and Mutation Taster,[p44] which accept a transcript reference, the affected coding position, or the type of mutation as input. Truncating variants in the first exons are typically prone to NMD, and an NMD cutoff in the coding sequence can be defined to select stable variants for further progression along the pipeline. Most of these variants are missense changes involving single amino acid substitutions. In such cases, conservation tools (e.g. ConSurf[p45]) are used to predict the significance of each residue in the protein structure, and thus to infer the likelihood that the change will affect protein function. Notably, from this point on, there is a need to consider the 3D structure of the protein responsible for the disease and of the related mutant proteins. When experimental protein structures are not available, AlphaFold models[p30],[p46],[p47] can be used as acceptable input.[p48] The 3D mutant proteins are modelled using the Chimera 'Rotamer' tool,[p49] which involves substituting the mutant amino acid within the 3D structure and optimizing the sidechain orientations.[p50] Optionally, more advanced tools may be used to refine the backbone structure.

The 3D structures of the protein variants are then compared with the physiological (wild type) counterpart to predict the effect of the mutation on molecular stability. To this end, dynamic profiles that include assessments of stability, flexibility, and interatomic interactions are compared. This is achieved using two techniques: normal vibrational mode analysis and molecular dynamics (with DynaMut2[p51] and CABSFlex,[p52]

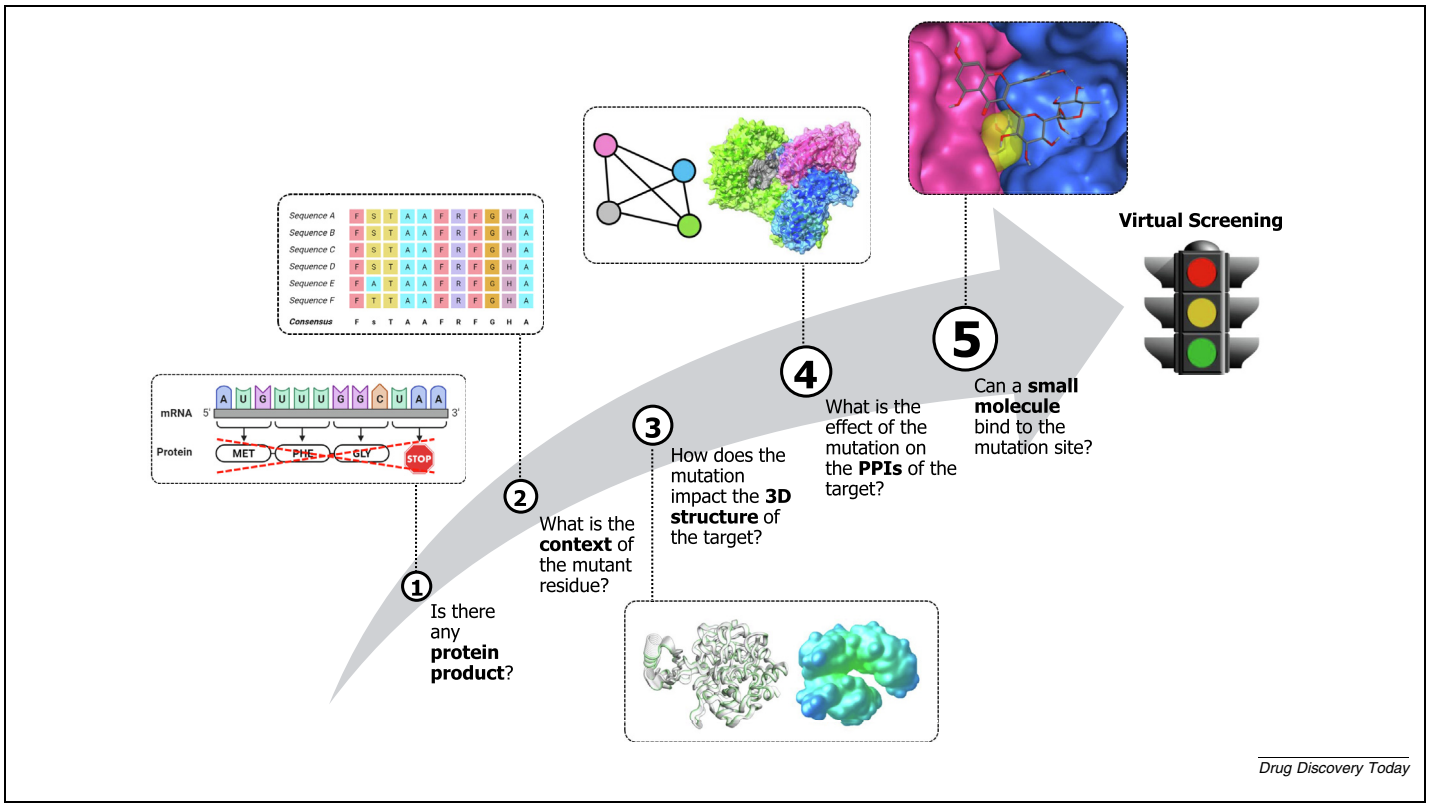

FIGURE 3

***In silico* steps required prior to drug repurposing.** PPI, Protein–protein interaction.

respectively). DynaMut provides both numerical and graphical outputs, enabling a quick evaluation of the mutation's effect, whereas CABSFlex 2.0 focuses on analyzing protein flexibility. Alternatively, machine learning models have acceptable performances as non-first-principles methods to achieve similar results.[p53]

Further insights are gained by comparing the molecular and solvent-accessible polar surface areas, which give an indication of the different interaction potentials of mutant proteins. Tools such as VEGAZZ[p54] and the Creighton method[p55] are used for this purpose. VEGAZZ allows for visual depiction of the variation, whereas the Creighton method estimates the relative solvent accessibility (RSA), which can be calculated using tools such as the free SASA package.[p56]

Another common pathogenic mechanism in human beings involves the disruption of protein–protein interactions (PPIs). This is particularly important for many proteins associated with HSP, which frequently perform their physiological functions within large protein complexes and which form oligomers. Intrinsically disordered regions (IDRs) often contain short recognition motifs for other proteins.[p57] Thus, pathogenic mutations that span IDRs are likely to disrupt such protein recognition motifs. To formulate hypotheses in this sense, a prediction of IDRs is obtained using IUPRED.[p58] However, PPIs also occur through structured regions of the involved proteins. To infer whether such interactions are affected, the known interactome is searched: databases (such as STRING[p59]) and pathway maps (e.g., from KEGG[p60] and Reactome[p61]) are queried before the PPIs are modelled by homology (using SwissModel and AlphaFold, for example). Mutant residues that are located at the interaction surfaces are identified as having the potential to affect the stability of PPIs. Overall, 3D structures enable the detailed characterization of the molecular features responsible for structural destabilization and/or alterations in PPIs.

The next step is to predict the druggability of the protein with small molecules. To this end, it is crucial to identify a binding pocket near the mutation site. Visual inspection of surface maps can be complemented by the use of both commercial and free software tools: (i) Site Finder,[p62] which is available in the commercial MOE modelling suite (https://www.chemcomp.com); (ii) fpocket,[p63] an open-source alternative for identifying binding pockets; and (iii) DoGSiteScorer,[p64] a free, server-based tool. The predictions made thus far provide semi-quantitative information about the feasibility of a drug-repurposing strategy. Their integration within a rational decision-making framework is illustrated in Figure 4. Next, docking is leveraged for structure-based virtual screenings (VS) to suggest ligands. The workflow involves curating a dataset of therapy-approved compounds (such as DrugBank[p65]) and organizing them in the appropriate input format. Docking is then performed using both commercial and free tools (such as Gold[p66] and Autodock Vina[p67]) through a consensus approach. Scoring the compounds and visualizing them within protein complexes allows for the identification of potential mutation-specific drug-repurposing strategies. Molecules

TABLE 2

Steps in a computational pipeline for drug repurposing and the relevant tools.

Step	Question	Criterion	Exemplar tools
1	Is there a protein to target?	Presence of a protein product downstream to the mutation	NMDesc, MutationTaster2021
2	What is context of the mutant residue?	Evolutionary conservation of the residue undergoing mutation	ConSurf
3	How does the mutation impact the 3D structure of the target?	a. Order/disorder prediction	IUPred2A
		b. Generation of mutant forms of the target	AlphaFold2/3, ESMFold, OmegaFold, OpenFOld, UCSF Chimera Rotamer tool
		c. Dynamic profile	DynaMut2, CABS-flex2
		d. Surface property variation	VegaZZ, freeSASA
4	What is the effect of the mutation on protein–protein interactions?	a. Evaluation of the interactome and pathway maps	STRING-DB, Reactome, KEGG
		b. Multimer modelling	SWISS-MODEL
5	Can a drug bind to the mutation site?	Presence of binding pockets	MOE-SiteFinder, fpocket, DoGSiteScorer
6	Is there any drug that can be repurposed?	Virtual screening	AutoDock Vina, GOLD

retrieved from the VS study are further manually screened to select those with the best therapeutic index and ADME profile, including blood–brain barrier (BBB) passage.

Experimental requirements for the validation of computational results in drug repurposing

Experiments are then required to validate virtual results and to complete the preclinical investigation. Various materials and instruments can be employed, allowing for different depths of experimental characterization. Not all of these processes are strictly required: the level of experimental characterization can be tuned depending on how much information is available in the scientific literature and how solid the computational results are. As a general rule, we suggest privileging proof-of-concept experiments first.

Protein characterization and the related biological toolkits

Characterizing the wild type protein is highly informative (Figure 5). This process includes assessing the intracellular localization of the protein and its organization as a monomer or in a homo- or heteromer. To achieve this, proteins are typically expressed as fusion constructs, often tagged or mutated. The availability of suitable plasmids for the transformation of these fusions is pivotal. A useful resource is the repository Addgene (www.addgene.org), which provides a wide range of commercial and proprietary plasmids. When investigating novel genes/mutations, however, it is likely that a custom-made plasmid must be cloned in-house or through contract research organizations. The next step involves transformation and protein expression. The selection of the optimal expression organism is crucial at this stage. Bacterial systems are generally faster and more cost-effective, but they have limitations such as their inability to preserve relevant-to-human post-translational modifications. On the other hand, eukaryotic cell lines typically result in lower protein yields. The decision about the expression system should be made by an expert within the team, typically a biotechnologist. Protein purification and biochemical analysis follow this step. Various isolation and purification procedures are employed, including chromatography (gel filtration, ion exchange, affinity, or hydrophobic interaction chromatography) and immunoprecipitation procedures.

Once pure or enriched protein fractions have been obtained, the size of the protein complex can be analyzed using gel filtration chromatography or density-gradient centrifugation. In addition, if the protein complex includes other proteins besides the target protein (i.e., is a heteromeric complex), it is necessary to identify the interacting and/or associated proteins. To achieve this, hypotheses can be formulated on the basis of information from interactome databases, and validated by co-immunoprecipitation in conjunction with analytical techniques (such as mass spectrometry). Depending on the subsequent experiments, a selective antibody might be necessary, and this can be chosen on the basis of information in the literature, in databanks such as antibodypedia[(p68)] and in catalogs from specific vendors. Technique-specific validation of the antibody (by immunostaining or western blotting) is necessary. Moreover, it is important to determine whether the epitope (i.e., the part of the protein that is recognized by the antibody) is known. This becomes particularly relevant when dealing with missense variants that may affect the ability of the antibody to recognize the protein.

Cell models

The availability of suitable cell models to monitor the disease is crucial for evaluating the effect of a drug. These models can be established from patient-derived tissues, obtained from colleagues, or acquired from repositories (such as the American Type Culture Collection [ATCC; https://www.atcc.org]). It is generally recommended that the simplest and most easily cultivable cell model is chosen, unless specific cell-type characteristics (such as neuronal morphology or electrophysiology) are to be investigated.

One effective *in vitro* strategy for studying a rare disease is the use of patient-derived skin fibroblast lines.[(p69),(p70)] Fibroblasts can be obtained from a skin punch biopsy, a minimally invasive

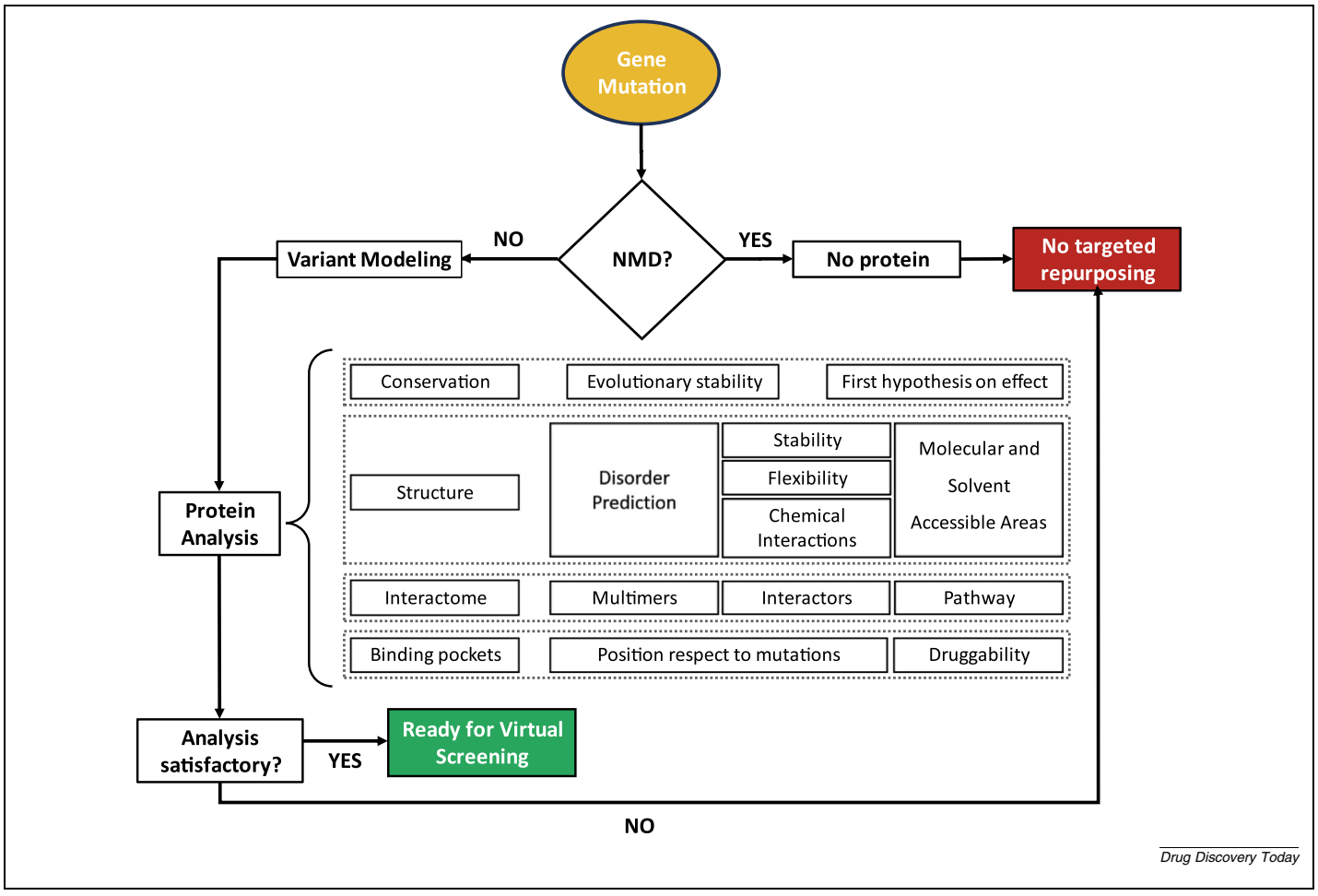

FIGURE 4

Selection criteria and workflow for assessing the druggability of a specific variant and determining the feasibility of virtual screening. NMD, Nonsense-mediated decay.

procedure, even from pediatric patients. This approach yields a robust and easily cultivable cell line, which is syngenetic to the patient and does not require genetic manipulation. Nevertheless, it is first important to determine whether the protein of interest is expressed in these cells. This can be addressed by consulting public repositories such as The Human Protein Atlas and GTEx, to compare expression levels in various tissues and cell lines. Furthermore, information on the intracellular localization can be obtained from public repositories (such as UniProt). Fibroblasts are a good model for studying proteins that are involved in common cellular mechanisms, such as mitochondrial dynamics and morphology. They can also be considered a viable model for investigating neurodegenerative changes because of their metabolic and biochemical relationships with neurons. Studies have shown that impairment of mitochondrial bioenergetics is found in fibroblasts from patients diagnosed with neurodegenerative pathologies such as Alzheimer's, Parkinson's and Huntington's diseases, as well as amyotrophic lateral sclerosis.[(p70),(p71)] Experiments can unveil differences in protein expression and localization by comparing fibroblasts from patients with healthy fibroblast lines, providing biochemical evidence of the impact of a specific variant. Nevertheless, it is important to consider that the genetic background of the healthy individuals from whom control fibroblasts are derived may mask underlying effects. To address this, it is advisable to use multiple control lines, possibly with age and sex matching.

Despite the advantages offered by fibroblasts, these tissues cannot be regarded as a comprehensive cell model for all situations. For instance, studying HSPs requires monitoring of neuron-specific features such as connectivity, morphology, and electrophysiology. In this case, fibroblasts can be reprogrammed to form induced pluripotent stem cells (iPSC). These iPSCs are subsequently differentiated into neurons,[(p72)] which allow accurate representation of specific mechanisms. Patient-derived neurons offer several advantages, including the ability to recapitulate complete phenotypes by maintaining the patient's genetic background. This provides a better understanding of disease mechanisms and a relevant model for studying neurodegenerative diseases. Nevertheless, the reprogramming and differentiation processes are time-consuming, require specialized expertise,

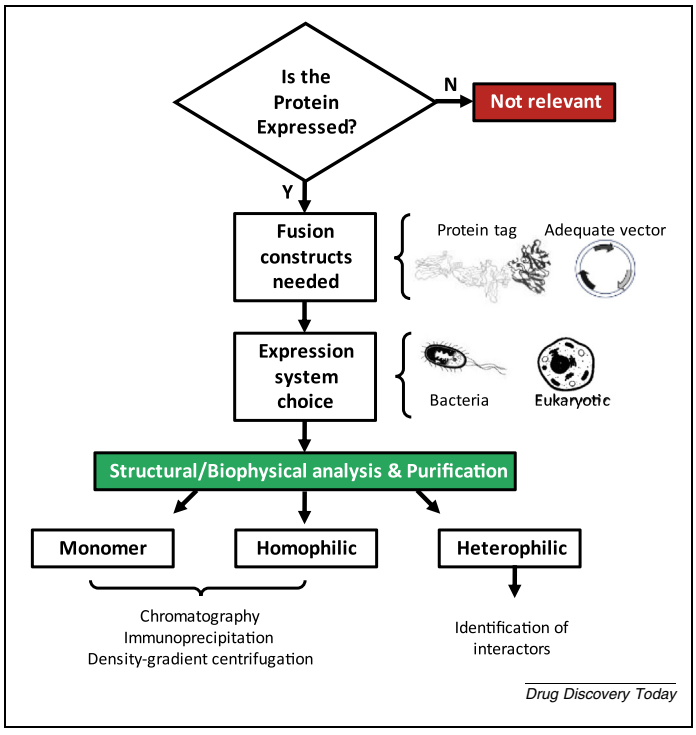

FIGURE 5

Flowchart outlining strategies for protein expression, protein purification, and the study of protein complexes.

and can be costly. As a result, this approach is well-suited for studying disease mechanisms and for the final testing of a drug candidate, but is likely to be too expensive to be used in a screening phase.

It is important to consider that the more representative a cell model is, the closer the study will align with the patient's reality. Thus, it is advisable to choose different systems carefully in light of the specific study goals balanced with the available time and resources. In practical terms, when implementing a drug-repurposing strategy for rare diseases, we advise that candidate drugs are initially prioritized using simple cell models, before more sophisticated systems are employed to validate final hypotheses and to further investigate the efficacy of potential treatments. With this approach, researchers can maximize the utility of their resources and make informed decisions.

Animal models

This manuscript aims to outline preclinical strategies for obtaining a drug candidate, and we primarily focus on experimental verification in cell-based systems. A comprehensive description of *in vivo* strategies is beyond the scope of this study, but we provide a brief commentary on what occurs beyond the *in vitro* phase to emphasize the need to consider the entire preclinical part of a drug discovery process. A thorough description of strategies employing animal models has been provided by some of the authors for the paradigmatic case of Rett syndrome.[p73] The availability of disease-specific mouse models is crucial, particularly in the context of gene therapy solutions. For many rare diseases, however, models are limited and often restricted to gene knockouts. Although valuable for studying gene functions, such models may not accurately represent the diversity of missense mutations. Furthermore, in conditions such as HSPs and other motor neuron diseases, mouse models exhibit milder or even absent phenotypes. This aspect should be carefully considered when planning a study, especially considering the FDA Modernization Act 2.0 (2022) which relaxes the requirement for animal model proof-of-concept in the approval of treatments (https://www.fda.gov). Along this line, future strategies include organ-on-chip approaches, which are intended as an adequate intermediate step. Furthermore, it has been proposed that holistic human-on-chip approaches have potential in the discovery of therapeutics for rare diseases.[p74]

Selected frontier tools: cryoEM, advanced microscopy techniques and electrophysiological characterization

Protein production, purification, and adequate cell models are requirements for validating virtual results. In many cases, these steps are sufficient to obtain compassionate use approval.[p11],[p75] However, further investigation might be required to gain a deeper understanding of the underlying biological mechanisms. Advanced tools and techniques, such as cryo-electron microscopy (cryo-EM), can confirm structural predictions, as can advanced microscopy techniques and electrophysiological characterization, which can also help in the identification of cell markers (Figure 6). By utilizing these frontier tools, researchers can also potentially identify novel targets for drug repurposing.

Knowledge of the 3D protein structure is crucial for a thorough understanding of protein function and the effects of mutations on protein stability and overall assembly. There have been recent advances in *de novo* protein structure prediction with the introduction of tools such as AlphaFold 2/3 and ESM-Fold combined with molecular dynamics. Nevertheless, experimental validation through experimental methods such as X-ray crystallography, NMR, and cryo-electron microscopy (cryoEM) remains essential in some cases.[p76] CryoEM is a particularly valuable technique due to its ability to offer high-resolution protein information and distinct advantages such as the ability to study proteins without the need for protein crystallization. These advantages significantly enhance the capability to analyze both membrane proteins and delicate, large protein complexes. By combining advanced sample preparation techniques with cutting-edge grid technologies (such as graphene and affinity grids), researchers can work with low volumes and protein concentrations approaching just 1 μg/ml. This makes it possible to access proteins that are difficult to overexpress or isolate from near-native conditions.[p77],[p78] Finally, cryoEM can capture snapshots of a protein in various conformational states (Figure 6A), enabling detailed exploration of the dynamics underlying a protein's molecular function. For multimeric proteins, for example, the structures of both the complex and the individual monomer are often unknown. By visualizing the protein's structure using cryoEM and subsequently studying mutant disease states, researchers can greatly enhance their understanding of the protein at the molecular level.

Cutting-edge microscopy encompasses a valuable set of techniques that can be employed to monitor pathologic phenotypes and to test drug candidates *in vitro* (Figure 6B). Confocal micro-

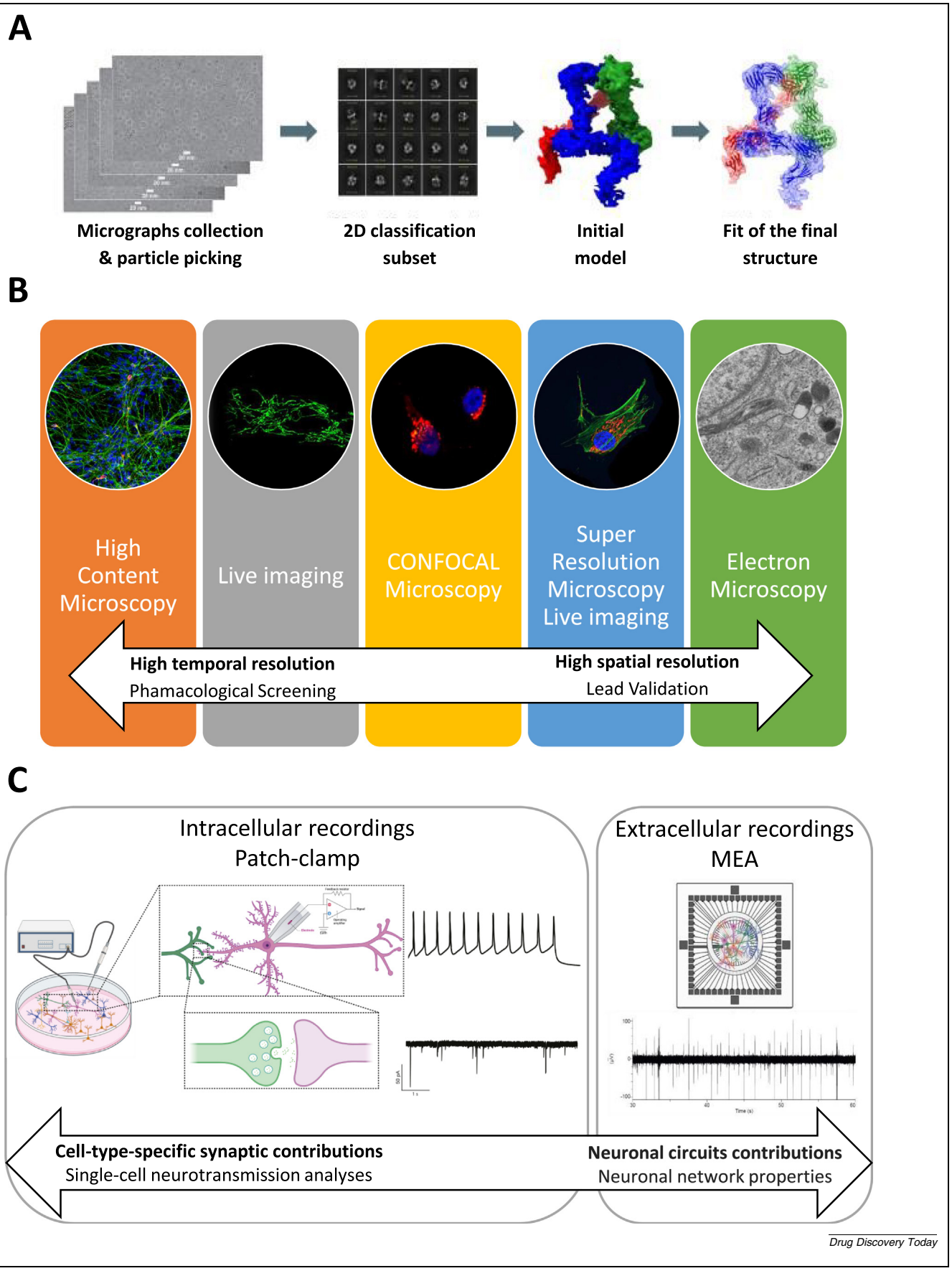

FIGURE 6

Selected frontier tools. A. Workflow for the determination of protein structure using cryoEM.[p78] The process involves the collection of CryoEM micrographs, particle picking, 2D classification and particle image averaging, 3D classification, and 3D map construction or refinement. The resulting map is used to fit a model of the full protein sequence. **B.** Advanced microscopy techniques: advantages and suggested application domain. **C.** Electrophysiological characterization.

scopy (CM) is the primary tool enabling the acquisition of high-resolution, 3D images of cellular structures. It is particularly valuable in studying neuronal elements such as axons and dendrites, and provides detailed insights into the impact of genetic mutations on cellular connectivity and function.[p79] Super-resolution microscopy (SRM), including structured illumination microscopy and stochastic optical reconstruction microscopy, goes beyond the diffraction limits of conventional light microscopy,[p80] allowing the visualization of subcellular structures at an unprecedented level of detail and significantly contributing to unraveling the structural alterations and molecular interactions associated with genetic mutations.[p81] Live-cell imaging (LI), notably time-lapse microscopy, offers dynamic insights into cellular processes and can be implemented in association with the primary techniques cited above. For HSP, LI has proven instrumental in tracking the real-time movement and behavior of cells and subcellular organelles, unraveling the dynamic impact of genetic mutations on cellular physiology. This information is paramount for assessing the efficacy of small molecules and gene therapies in reinstating normal cellular function. Finally, high-content screening microscopy currently integrates CM, SRM and LI, enhancing the role of microscopy in validating drug discovery through the concurrent assessment of multiple cellular parameters.[p82] Transmission electron microscopy, although not conventionally considered a high-throughput technique, has emerged as a valuable addition to the microscopy arsenal, unveiling ultrastructural details at the nanoscale.[p83]

When looking for cellular markers, electrophysiological assays can provide information on both physiological and pathological properties (such as spontaneous and evoked activity, synaptic transmission and plasticity), as well as on the efficacy of treatments (Figure 6C). Both intracellular (i.e., patch clamp) and extracellular (i.e., multielectrode array [MEA] and field potential) recordings are robust investigational strategies to assess how altered neuronal responses underlie specific behaviors or clinical manifestations.[p84] Intracellular recordings performed with patch clamp, in particular on isolated cells in culture, can precisely define the complex biophysical characteristics of ion channels and neuronal excitability, fundamentally contributing to the comprehension of the neurophysiological bases of rare diseases.[p85] The patch clamp technique allows characterization of the impact of drugs on single ion channel function, whereas the multicomponent properties of either intact or reconstructed neuronal networks can be evaluated by MEA analyses. The MEA approach has advantages over single-cell recordings for high-throughput drug screenings,[p84] allowing the generation of multi-parametric, high-resolution data to profile the drug impact in longitudinal settings on electrically active cells (such as muscle and neural cells). Patients who are affected by HSP have been tested with non-invasive procedures such as electromyography and measurements of motor evoked potentials, maximal nerve conduction velocity, and compound muscle action potentials. This work revealed severe dysfunction of the corticospinal tracts caused by the degeneration of upper motor neurons.[p86] These *in vivo* analyses point out the importance of understanding the electrophysiological bases of upper motor neuron dysfunction in HSP cellular models by applying either single-cell or network electrophysiology assays. This aim can be achieved by studying the motor cortex of the available HSP murine mouse models or cortical motor neurons derived from patient iPSCs.

Overall, Cryo-EM, cutting-edge microscopy, and electrophysiological analyses are expected to provide a detailed understanding of the structural, morphological, and physiological aspects of rare diseases. As an example, current investigations on the ultra-rare infantile ascending hereditary spastic paralysis, which is caused by defects in the Alsin protein,[p87] are ongoing in our laboratories. Carrying out such experiments is not a trivial task, however, with target-specific issues requiring expert troubleshooting. For instance, buffer compatibility and protein stability affect the resolution of cryoEM, and variability in cell cultivation protocols can impact microscopy and electrophysiology experiments.

Common issues that rare disease drug candidates must overcome

In the development of drug candidates for rare diseases, particular focus should be put on answering fundamental questions relating to common issues in drug discovery. We briefly discuss them below.

Administration route and the blood–brain barrier permeability

The administration route is the first crucial factor that determines the ability of a drug to reach its active site effectively.[p88] This work does not delve into the specifics of all possible administration routes, but it is important to note that the administration route is especially relevant for medications targeting the central nervous system (CNS) because of the presence of the blood–brain barrier (BBB). In general, direct administration into the brain parenchyma, using invasive techniques such as intrathecal administration, is the preferred route for large, single-dose drugs.[p35] For small molecules, the oral or parenteral route is typically chosen, and it is essential to ensure that the drug candidate can cross the BBB.

The range of computational methods available for predicting and determining the passage of drugs into the CNS (expressed as log BB, a logarithmic ratio of the concentrations of drug in the brain and the blood) has significantly expanded.[p89],[p90] Nevertheless, these models often show limited predictive power for molecules outside the class on which the model was originally trained. Moreover, efflux transporters, such as P-glycoprotein (P-gp) and many others, actively remove xenobiotics from the CNS. This complexity is often not considered by the most common predictors.

Cell-based models using brain endothelium have also been developed and validated to access BBB permeability experimentally.[p91] Some of these are suitable for medium to high throughput screening, but several limitations are present. The source material (primarily from mouse or rat or from differentiated from pluripotent cells) comes with high maintenance costs. Conversely, cheaper models based on immortalized cell lines are less representative and less reproducible.[p92] Overall, methods for evaluating the BBB passage of new drug candidates are available, particularly for small molecules, but their reliability is not optimal.

Biomarkers to monitor the efficacy of the treatment

Testing new experimental therapies and building more robust and better stratified clinical trials requires biomarkers that correlate with relevant clinical endpoints and are easily measurable over time. Descriptions of the outcome measures and of useful biomarkers in the clinical HSP context are beyond the scope of this paper and can be found elsewhere.[p93] However, we recall that the term 'biomarker' or 'biological marker' refers to a broad subcategory of medical signs that can be measured accurately and reproducibly. As defined by the US Food and Drug Administration, a biomarker is "a characteristic that is measured as an indicator of normal biological processes, pathogenic processes, or responses to an exposure or intervention". Biomarkers might include clinical scales, neuroimaging features, and changes in body fluids, and can be further classified according to their use. Examples include biomarkers that can be used for diagnostics, monitoring, response prediction, prognosis, and risk susceptibility.[p94] Among these, fluid biomarkers are the least invasive, least expensive, and easiest to obtain. These biomarkers could be derived from any bodily fluid, including blood, cerebrospinal fluid, urine, and saliva.

At the preclinical stage and when a biological marker has not yet been established, researchers look for cell markers: proteins, metabolites, or specific phenotypes that are characteristic of the disease model under study. Some features of neurodegenerative diseases can be used for this purpose;[p95] for instance, mitochondrial impairment,[p96] decreased ATP production, excessive generation of reactive oxygen species, and dysregulation of calcium levels are cell markers for some HSPs. In practice, monitoring variation in cell markers may help both in identifying the modifications induced by the variants and in determining the effect of a drug candidate that has been identified by virtual screening.

Techniques such as omics approaches (such as transcriptomics, proteomics and metabolomics) hold potential to study differential gene expression and the impact of specific mutations in disease- or patient-representative cell lines. In addition, more advanced methods (such as pathway mapping) will help to reveal which signaling networks are affected. Nevertheless, it is important to note that such approaches are limited to assessing just transcriptional regulation.[p97] Further information comes from proteomic studies, a group of techniques characterizing the translated proteome in both qualitative and quantitative terms. Protein-level approaches yield more mechanistic insights, but in contrast to transcriptomic approaches, they do not yet allow investigation of the whole proteome with a single untargeted assay.[p98] Finally, the study of metabolites holds promise to study the effect of a gene mutation by analyzing changes in metabolite profiles, identifying cell markers and providing the basics to identify biomarkers.[p99] However, the dynamic nature of metabolites represents a limitation, often resulting in extreme variability.[p99]

Underestimation of risk is a key project weakness

Alternative approaches for drug discovery are necessary for rare diseases. This requires a rational prioritization of the efforts and careful planning. It is not uncommon for specialists within the team to underestimate the time and resources required to complete key experiments. To address this challenge, the team leader must identify and consider the hurdles associated with each critical step of the project. Examples might include computational artifacts, issues with protein purification or stability, lack of protein expression in cell models, the size of the genes for gene therapy, and the toxic gain of function by aberrant protein products.

Case studies

Below, we discuss two recent successful drug discovery efforts for two monogenic rare HSP diseases, each based on a different approach.

Infantile-onset ascending hereditary spastic paralysis (IAHSP): a small molecule success

Infantile onset hereditary spastic paralysis (IAHSP) is an ultra-rare neurological disorder that affects fewer than 100 children worldwide.[p100] This condition is caused by biallelic pathogenic autosomic variants in the *ALS2* gene. This gene encodes for Alsin, which plays a critical role in the differentiation and maintenance of upper motor neurons.[p101] The gene is approximately 95.5 kb in size, with a cDNA sequence of 6,675 bp that encodes the 1,657 amino acid protein. *ALS2* is larger than the upper limit for translocation in adeno-associated viruses (AAVs), which are commonly used as gene therapy vectors.[p102] Therefore, gene therapy strategies for IAHSP are premature, with actual technologies and small molecules preferable. Treatment with mutation-specific small molecules is only possible for missense, non-destabilizing variants. Analysis of IAHSP patient databases revealed that some individuals have missense changes that destabilize the overall protein structure to a minor extent.[p48] This finding confirms that targeting these specific mutations with small molecules may be a viable therapeutic strategy for these patients.

A recent drug repurposing effort identified a potential drug candidate, menatetrenone (also called MK4, CAS number 863-61-6), for the treatment of a specific *ALS2* variant leading to IAHSP[p11] (Figure S1A). MK4 was found to bind and restore the function of Alsin with the missense variant R1611W, which was originally discovered in an Italian patient.[p11] This patient displays a compound heterozygous genotype, with the other mutation being a frameshift change that is associated with loss-of function. Thus, menatetrenone can only act on the protein with the R1611W change. Nevertheless, the fact that the heterozygous parents are clinically indistinguishable from healthy individuals provides evidence that just one functional copy of the gene is enough to prevent any symptoms. The identification of MK4 as a potential treatment for IAHSP is a significant milestone in our research as this discovery was made using *in silico* techniques, later validated with selected *in vitro* experiments.[p11] Our results led to the approval for compassionate use of a patient-specific therapeutic regimen. Such approvals are granted by medical committees for already-approved molecules when no alternative cure is available. Currently, the patient is taking MK4 orally on a thrice-daily basis with promising results. This case serves as a paradigm for a drug-repurposing strategy achieved through an alternative drug discovery model.[p11] The research and development efforts were supported by a patient's association, HelpOlly (www.helpolly.it, Figure 2),

highlighting the collaborative nature of this work and the importance of patient advocacy.

In hindsight, we identified the key pieces of knowledge that were crucial for the successful repurposing of MK4. First, prior biochemical research conducted by some of us showed that Alsin exists as an active tetramer in the cytosolic compartment.[p103] This information was instrumental in rationalizing the pathogenesis of the R1611W variant: 3D protein modeling predicts that the mutant version of Alsin prevents the formation of tetramers by inducing abnormal dimers (Figure S1B).[p11] Through virtual screening, we identified MK4 as a small molecule that masks the mutant residue and restores proper tetramerization. Second, prior knowledge about the involvement of Alsin with mitochondria[p104] suggested that we should monitor mitochondrial morphology using advanced microscopy techniques in patient-derived fibroblasts, which allowed the identification of a cell marker and validated the use of MK4 (data to be published).

This result satisfies two important criteria for treating rare diseases: it is highly personalized and it requires minimal resources. However, MK4 is only effective for the R1611W mutation and multiple other missense mutations are known. This means that similar studies must be conducted for each mutation, with no guarantee that suitable therapeutic molecules will be found for each case. On the positive side, having a successful pilot study provides the necessary biological tools and allows for efficient testing of other molecules.

SPG50: a gene therapy success

Spastic paraplegia type 50 (SPG50) is an extremely rare neurodegenerative disease with an autosomal recessive inheritance pattern (OMIM #612936). At present, there are only about 80 reported cases of SPG50 worldwide, originating from loss-of-function mutations in the *AP4M1* gene.[p32] The gene encodes the medium subunit of the adaptor protein complex 4 (AP-4), which is involved in the intracellular vesicular trafficking of transmembrane proteins.[p105] Despite much knowledge of the cell biology of the disease, no specific treatment is currently available for SPG50, and the pursuit of a drug repurposing strategy is complicated by the lack of a 3D experimental structure. Some attempts have been made to model key interactions and to provide preliminary structures of the AP4 complex.[p106] Starting from these, we are currently investigating specific missense mutations that are associated with SPG50, with the two-fold aim of rationalizing their pathogenesis and investigating patient-specific drug repurposing opportunities.

The relatively small size of *AP4M1* supports gene therapy as a feasible strategy to treat SPG50, and other groups are investigating avenues that do not involve small molecules. In collaboration with the nonprofit association CureSPG50 (www.curespg50.org; Figure 2), Steven Gray's lab has contributed to the development of MELPIDA, an AAV9-mediated gene therapy for SPG50[p35] that is currently in Phase II clinical trials. The therapy involves delivering a functional copy of the human *AP4M1* cDNA via an AAV9 vector, using intrathecal injection to bypass the BBB passage. Promising results have been observed in preclinical *in vitro* and *in vivo* studies, with minimal adverse effects in *AP4M1*-deficient models.[p35] The first patients have already been treated with encouraging outcomes, and the clinical trials will provide a definitive answer regarding its efficacy. The estimated primary completion date for the trials is October 1, 2028, with the study completion set for October 1, 2030. The approach used for MELPIDA has the potential to be applied to all AP4M1 mutants, but obtaining approval for a new treatment, especially one that utilizes cutting-edge technologies such as AAV-associated gene delivery, requires considerable financial resources. For this reason, MELPIDA showcases the importance of patients' associations in driving the development of treatments. As for the case of IAHSP, the central role of associations such as CureSPG50 in securing adequate funding, and the involvement of the patients' families in coordinating efforts, highlights the value of alternative drug discovery models for addressing rare diseases.

Summary and conclusions

The pharmaceutical industry tends to prioritize common conditions, leaving the needs of patients with rare diseases largely unmet. As a result, smaller and more focused entities such as patients' associations, nonprofit foundations, and families often take the initiative. They play a crucial role in securing financial support for research projects through activities like knowledge dissemination and fundraising. However, such resources are typically small when compared to the budgets allocated for traditional drug discovery paths. Therefore, finding new treatment options for rare diseases requires alternative models.

In this review, we used a group of rare monogenic diseases, including HSPs, as examples to provide guidelines for the design and preclinical efforts necessary to discover drugs. After discussing and analyzing the composition of the ideal research team for a specific project, our attention shifted towards the necessary data collection. This involves creating a comprehensive dataset that is specifically designed to collect patient clinical and genetic data, as well as protein-related information from various sources.

Subsequently, we explored the potential paths for developing treatments for rare monogenic diseases such as HSPs, focusing on two main approaches: gene therapy and small molecules. These strategies have different areas of application and varying costs of development, but they are not always mutually exclusive. When choosing small molecules as a strategy, we, along with other colleagues, propose that drug repurposing brings significant benefits. We carefully assessed the role and impact of computational strategies in identifying cases where drug repurposing is feasible, beyond the exclusive use of docking procedures. We emphasize that modern predictors of 3D protein structures provide valuable information for this purpose. Although there is an abundance of computational tools, most of which are freely available, they are not always validated and lack rational organization, making their routine application challenging. Therefore, we propose an analysis flow for each step of the process.

To validate the computational predictions at the preclinical level, essential experimental tools are required for efficient testing with the aim of providing readouts in the most efficient manner. Specifically, we focused on protein expression and cell models, which are crucial for proof-of-concept experiments to validate computationally identified drug-repurposing candidates. Furthermore, we discussed advanced tools that, although not

strictly essential, can greatly enhance our understanding of the molecular and cellular mechanisms underlying the disease. Examples are cryo-electron microscopy, advanced microscopy techniques, and electrophysiological functional assays.

Finally, we presented two case studies that required different approaches. First, a IAHSP patient with a missense variant in the *ALS2* gene, leading to a stable protein that can be targeted with repurposed small molecules. Second, a case of SPG50 in which the delivery of the entire *AP4M1* cDNA via an AAV9 vector is necessary. As general concept, gene therapy can be seen as a curative approach, but it requires novel approvals and entails significant expenses. On the other hand, drug repurposing targets specific mutations, making it applicable only to specific patients.

In summary, we suggest that, given the often low budgets for rare diseases, a drug repurposing strategy based on the pipeline presented in this review holds significant potential and should be pursued whenever possible.

Additional study group contributors in the IAHSP Consortium

Giulia Caron, University of Torino, Molecular Biotechnology and Health Sciences Department, CASSMedChem, 10138 Torino, Italy. **Lorenzo Lamacchia**, University of Torino, Molecular Biotechnology and Health Sciences Department, CASSMedChem, 10138 Torino, Italy. **Serena Francisco**, University of Torino, Molecular Biotechnology and Health Sciences Department, CASSMedChem, 10138 Torino, Italy. **Alfredo Brusco**, Department of Neurosciences Rita Levi-Montalcini, University of Turin, Turin, Italy; Medical Genetics Unit, Città della Salute e della Scienza University Hospital, Turin, Italy. **Verdiana Pullano**, Department of Medical Sciences, University of Torino, Turin, Italy. **Gabriele Baj**, Department of Life Sciences, University of Trieste, Trieste, Italy. **Antonia Gurgone**, Department of Neurosciences Rita Levi-Montalcini, University of Turin, Turin, Italy. **Giuseppe Chiantia**, Department of Neurosciences Rita Levi-Montalcini, University of Turin, Turin, Italy. **Maurizio Giustetto**, Department of Neurosciences Rita Levi-Montalcini, University of Turin, Turin, Italy. **Silvia Russo**, Research Laboratory of Medical Cytogenetics and Molecular Genetics, IRCCS Istituto Auxologico Italiano, Milan, Italy. **Anna J. Higgins**, PUXANO, Ottergemsesteenweg Zuid 713, 9000 Gent, Belgium. **Wouter van Putte**, PUXANO, Ottergemsesteenweg Zuid 713, 9000 Gent, Belgium. **Stefania Della Vecchia**, Department of Developmental Neuroscience, IRCCS Stella Maris, Calambrone, Pisa, Italy; IRCCS Stella Maris Foundation, Molecular Medicine for Neurodegenerative and Neuromuscular Disease Unit, Pisa, Italy. **Roberta Battini**, Department of Developmental Neuroscience, IRCCS Stella Maris, Calambrone, Pisa, Italy; Department of Clinical and Experimental Medicine, Neurological Clinic, University of Pisa, Pisa, Italy. **Filippo Maria Santorelli**, IRCCS Stella Maris Foundation, Molecular Medicine for Neurodegenerative and Neuromuscular Disease Unit, Pisa, Italy.

CRediT authorship contribution statement

Matteo Rossi Sebastiano: Writing – review & editing, Writing – original draft, Visualization, Investigation, Data curation, Conceptualization. **Shinji Hadano:** Writing – review & editing, Writing – original draft, Visualization. **Fabrizia Cesca:** Writing – review & editing, Writing – original draft, Visualization. **IAHSP Consortium:** Writing – review & editing, Writing – original draft, Visualization, Investigation. **Giuseppe Ermondi:** Writing – review & editing, Writing – original draft, Visualization, Conceptualization.

Funding

FMS and RB are supported in part by the Italian Ministry of Health-Ricerca Corrente 2024. F.C. and GE are supported by Call PRIN 2022 - Prot. 20224YX5ZX - Investigating common pathogenic mechanisms of rare genetic hereditary spastic paraplegia - CUP J53D23010800006 - Funded by European Union – Next Generation EU (M4.C2.1.1).

Data availability

No data was used for the research described in the article.

Acknowledgements

The Authors would like to thank the nonprofit associations HelpOlly (https://helpolly.it/en/) and CureSPG50 (www.curespg50.org/), which are dedicated to finding treatments for IAHSP and SPG50, respectively.

Conflicts of interest

The authors have no conflicting interests to disclose.

Appendix A. Supplementary material

Supplementary material to this article can be found online at https://doi.org/10.1016/j.drudis.2024.104138.

References

1. Delaye J, Cacciatore P, Kole A. Valuing the "burden" and impact of rare diseases: a scoping review. *Front Pharmacol*. 2022;13, 914338. https://doi.org/10.3389/fphar.2022.914338.
2. Marwaha S, Knowles JW, Ashley EA. A guide for the diagnosis of rare and undiagnosed disease: beyond the exome. *Genome Med*. 2022;14:23. https://doi.org/10.1186/s13073-022-01026-w.
3. Wang G et al. Rare and undiagnosed diseases: from disease-causing gene identification to mechanism elucidation. *Fundam Res*. 2002;2:918–928. https://doi.org/10.1016/j.fmre.2022.09.002.
4. Sakate R, Kimura T. Drug repositioning trends in rare and intractable diseases. *Drug Discov Today*. 2022;27:1789–1795. https://doi.org/10.1016/j.drudis.2022.01.013.
5. Alves VM et al. Knowledge-based approaches to drug discovery for rare diseases. *Drug Discov Today*. 2022;27:490–502. https://doi.org/10.1016/j.drudis.2021.10.014.
6. Wood J, Sames L, Moore A, Ekins S. Multifaceted roles of ultra-rare and rare disease patients/parents in drug discovery. *Drug Discov Today*. 2013;18:1043–1051. https://doi.org/10.1016/j.drudis.2013.08.006.
7. Shribman S, Reid E, Crosby AH, Houlden H, Warner TT. Hereditary spastic paraplegia: from diagnosis to emerging therapeutic approaches. *Lancet Neurol*. 2019;18:1136–1146. https://doi.org/10.1016/S1474-4422(19)30235-2.
8. Travaglini L et al. The impact of next-generation sequencing on the diagnosis of pediatric-onset hereditary spastic paraplegias: new genotype-phenotype

correlations for rare HSP-related genes. *Neurogenetics*. 2018;19:111–121. https://doi.org/10.1007/s10048-018-0545-9.
9. Blackstone C, O'Kane CJ, Reid E. Hereditary spastic paraplegias: membrane traffic and the motor pathway. *Nat Rev Neurosci*. 2011;12:31–42. https://doi.org/10.1038/nrn2946.
10. Pogue RE et al. Rare genetic diseases: update on diagnosis, treatment and online resources. *Drug Discov Today*. 2018;23:187–195. https://doi.org/10.1016/j.drudis.2017.11.002.
11. Rossi Sebastiano M, Ermondi G, Sato K, Otomo A, Hadano S, Caron G. Personalized treatment for infantile ascending hereditary spastic paralysis based on in silico strategies. *Molecules*. 2022;27:7063. https://doi.org/10.3390/molecules27207063.
12. Kibbelaar RE et al. Bridging the gap between the randomised clinical trial world and the real world by combination of population-based registry and electronic health record data: a case study in haemato-oncology. *Eur J Cancer*. 2017;86:178–185. https://doi.org/10.1016/j.ejca.2017.09.007.
13. Gliklich RE, Dreyer NA, Leavy MB, eds. *Registries for evaluating patient outcomes: a user's guide. Third edition*. Rockville (MD): Agency for Healthcare Research and Quality (US); 2014. Report No.: 13(14)-EHC111.
14. de Groot S et al. Balancing the optimal and the feasible: a practical guide for setting up patient registries for the collection of real-world data for health care decision making based on Dutch experiences. *Value Health*. 2017;20:627–636. https://doi.org/10.1016/j.jval.2016.02.007.
15. Khozin S, Blumenthal GM, Pazdur R. Real-world data for clinical evidence generation in oncology. *J Natl Cancer Inst*. 2017;109. https://doi.org/10.1093/jnci/djx187.
16. 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*. 2015;17:405–424. https://doi.org/10.1038/gim.2015.30.
17. Landrum MJ et al. ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res*. 2018;46:D1062–D1067. https://doi.org/10.1093/nar/gkx1153.
18. Fokkema IFAC et al. The LOVD3 platform: efficient genome-wide sharing of genetic variants. *Eur J Hum Genet*. 2021;29:1796–1803. https://doi.org/10.1038/s41431-021-00959-x.
19. Stenson PD et al. The Human Gene Mutation Database (HGMD®): optimizing its use in a clinical diagnostic or research setting. *Hum Genet*. 2020;139:1197–1207. https://doi.org/10.1007/s00439-020-02199-3.
20. Karczewski KJ et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. 2020;581:434–443. https://doi.org/10.1038/s41586-020-2308-7.
21. Abou Tayoun AN et al. Recommendations for interpreting the loss of function PVS1 ACMG/AMP variant criterion. *Hum Mutat*. 2018;39:1517–1524. https://doi.org/10.1002/humu.23626.
22. Adzhubei IA et al. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010;7:248–249. https://doi.org/10.1038/nmeth0410-248.
23. Ng PC, Henikoff S. SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res*. 2003;31:3812–3814. https://doi.org/10.1093/nar/gkg509.
24. Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res*. 2019;47:D886–D894. https://doi.org/10.1093/nar/gky1016.
25. Pertea M, Lin X, Salzberg SL. GeneSplicer: a new computational method for splice site prediction. *Nucleic Acids Res*. 2001;29:1185–1190. https://doi.org/10.1093/nar/29.5.1185.
26. Reese MG, Eeckman FH, Kulp D, Haussler D. Improved splice site detection in Genie. *J Comput Biol*. 1997;4:311–323. https://doi.org/10.1089/cmb.1997.4.311.
27. Kopanos C et al. VarSome: the human genomic variant search engine. *Bioinformatics*. 2019;35:1978–1980. https://doi.org/10.1093/bioinformatics/bty897.
28. Rodrigues EDS et al. Variant-level matching for diagnosis and discovery: challenges and opportunities. *Hum Mutat*. 2022;43:782–790. https://doi.org/10.1002/humu.24359.
29. The UniProt Consortium. UniProt: the universal protein knowledgebase. *Nucleic Acids Res*. 2017;45:D158–D169. https://doi.org/10.1093/nar/gkw1099.
30. Jumper J et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596:583–589. https://doi.org/10.1038/s41586-021-03819-2.
31. Racis L, Tessa A, Pugliatti M, Storti E, Agnetti V, Santorelli FM. Infantile-onset ascending hereditary spastic paralysis: a case report and brief literature review. *Eur J Paediatr Neurol*. 2014;18:235–239. https://doi.org/10.1016/j.ejpn.2013.09.009.
32. Jordan C, Geisel G, Alecu JE, Zhang B, Sahin M, Ebrahimi-Fakhari D. Disease severity and motor impairment correlate with health-related quality of life in AP-4-associated hereditary spastic paraplegia. *Neurol Genet*. 2021;7:e605.
33. Gunning AC et al. Assessing performance of pathogenicity predictors using clinically relevant variant datasets. *J Med Genet*. 2021;58:547–555. https://doi.org/10.1136/jmedgenet-2020-107003.
34. Korn D et al. Defining clinical outcome pathways. *Drug Discov Today*. 2022;27:1671–1678. https://doi.org/10.1016/j.drudis.2022.02.008.
35. Chen X et al. Intrathecal AAV9/AP4M1 gene therapy for hereditary spastic paraplegia 50 shows safety and efficacy in preclinical studies. *J Clin Investigation*. 2023;133:e164575.
36. Trapani I, Tornabene P, Auricchio A. Large gene delivery to the retina with AAV vectors: are we there yet? *Gene Ther*. 2021;28:220–222. https://doi.org/10.1038/s41434-020-0174-4.
37. Roessler HI, Knoers NVAM, van Haelst MM, van Haaften G. Drug repurposing for rare diseases. *Trends Pharmacol Sci*. 2021;42:255–267. https://doi.org/10.1016/j.tips.2021.01.003.
38. Nosengo N. Can you teach old drugs new tricks? *Nature*. 2016;534:314–336. https://doi.org/10.1038/534314a.
39. Van Norman GA. Phase II trials in drug development and adaptive trial design. *JACC Basic Transl Sci*. 2019;4:428–437. https://doi.org/10.1016/j.jacbts.2019.02.005.
40. Defelippe VM et al. Toward responsible clinical n-of-1 strategies for rare diseases. *Drug Discov Today*. 2023;28, 103688. https://doi.org/10.1016/j.drudis.2023.103688.
41. Hechtelt Jonker A et al. IRDiRC drug repurposing guidebook: making better use of existing drugs to tackle rare diseases. *Nat Rev Drug Discov*. 2023;22:937–938. https://doi.org/10.1038/d41573-023-00168-9.
42. Baker KE, Parker R. Nonsense-mediated mRNA decay: terminating erroneous gene expression. *Curr Opin Cell Biol*. 2004;16:293–299. https://doi.org/10.1016/j.ceb.2004.03.003.
43. Coban-Akdemir Z et al. Identifying genes whose mutant transcripts cause dominant disease traits by potential gain-of-function alleles. *Am J Human Genet*. 2018;103:171–187. https://doi.org/10.1016/j.ajhg.2018.06.009.
44. Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat Methods*. 2014;11:361–362. https://doi.org/10.1038/nmeth.2890.
45. Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal N. ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res*. 2010;38:W529–W533. https://doi.org/10.1093/nar/gkq399.
46. Tunyasuvunakool K et al. Highly accurate protein structure prediction for the human proteome. *Nature*. 2021;596:590–596. https://doi.org/10.1038/s41586-021-03828-1.
47. Abramson J et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*. 2024;630:493–500. https://doi.org/10.1038/s41586-024-07487-w.
48. Rossi Sebastiano M, Ermondi G, Hadano S, Caron G. AI-based protein structure databases have the potential to accelerate rare diseases research: AlphaFoldDB and the case of IAHSP/Alsin. *Drug Discov Today*. 2022;27:1652–1660. https://doi.org/10.1016/j.drudis.2021.12.018.
49. Pettersen EF et al. UCSF Chimera – a visualization system for exploratory research and analysis. *J Comput Chem*. 2004;25:1605–1612. https://doi.org/10.1002/jcc.20084.
50. Shapovalov MV, Dunbrack RL. A smoothed backbone-dependent rotamer library for proteins derived from adaptive kernel density estimates and regressions. *Structure*. 2011;19:844–858. https://doi.org/10.1016/j.str.2011.03.019.
51. Rodrigues CHM, Pires DEV, Ascher DB. DynaMut: predicting the impact of mutations on protein conformation, flexibility and stability. *Nucleic Acids Res*. 2018;46:W350–W355. https://doi.org/10.1093/nar/gky300.
52. Kuriata A et al. CABS-flex 2.0: a web server for fast simulations of flexibility of protein structures. *Nucleic Acids Res*. 2018;46:W338–W343. https://doi.org/10.1093/nar/gky356.
53. Caldararu O, Blundell TL, Kepp KP. Three simple properties explain protein stability change upon mutation. *J Chem Inf Model*. 2021;61:1981–1988. https://doi.org/10.1021/acs.jcim.1c00201.

54. Pedretti A, Mazzolari A, Gervasoni S, Fumagalli L, Vistoli G. The VEGA suite of programs: an versatile platform for cheminformatics and drug design projects. *Bioinformatics*. 2021;37:1174–1175. https://doi.org/10.1093/bioinformatics/btaa774.
55. Hiatt SM et al. Deleterious, protein-altering variants in the transcriptional coregulator ZMYM3 in 27 individuals with a neurodevelopmental delay phenotype. *Am J Hum Genet*. 2023;110:215–227. https://doi.org/10.1016/j.ajhg.2022.12.007.
56. Mitternacht S. FreeSASA: an open source C library for solvent accessible surface area calculations. *F1000Res*. 2016;5:189. https://doi.org/10.12688/f1000research.7931.1.
57. Dinkel H et al. The eukaryotic linear motif resource ELM: 10 years and counting. *Nucleic Acids Res*. 2014;42:D259–D266. https://doi.org/10.1093/nar/gkt1047.
58. Dosztányi Z. Prediction of protein disorder based on IUPred. *Protein Sci*. 2018;27:331–340. https://doi.org/10.1002/pro.3334.
59. Franceschini A et al. STRING v9.1: protein–protein interaction networks, with increased coverage and integration. *Nucleic Acids Res*. 2013;41:D808–D815. https://doi.org/10.1093/nar/gks1094.
60. Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000;28:27–30. https://doi.org/10.1093/nar/28.1.27.
61. Croft D et al. Reactome: a database of reactions, pathways and biological processes. *Nucleic Acids Res*. 2011;39:D691–D697. https://doi.org/10.1093/nar/gkq1018.
62. Krivák R, Hoksza D. Improving protein-ligand binding site prediction accuracy by classification of inner pocket points using local features. *J Cheminform*. 2015;7:12. https://doi.org/10.1186/s13321-015-0059-5.
63. Le Guilloux V, Schmidtke P, Tuffery P. Fpocket: an open source platform for ligand pocket detection. *BMC Bioinformatics*. 2009;10:168. https://doi.org/10.1186/1471-2105-10-168.
64. Volkamer A, Kuhn D, Rippmann F, Rarey M. DoGSiteScorer: a web server for automatic binding site prediction, analysis and druggability assessment. *Bioinformatics*. 2012;28:2074–2205. https://doi.org/10.1093/bioinformatics/bts310.
65. Wishart DS et al. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res*. 2018;46:D1074–D1082. https://doi.org/10.1093/nar/gkx1037.
66. Verdonk ML, Cole JC, Hartshorn MJ, Murray CW, Taylor RD. Improved protein-ligand docking using GOLD. *Proteins*. 2003;52:609–623. https://doi.org/10.1002/prot.10465.
67. Kochnev Y, Hellemann E, Cassidy KC, Durrant JD. Webina: an open-source library and web app that runs AutoDock Vina entirely in the web browser. *Bioinformatics*. 2020;36:4513–4515. https://doi.org/10.1093/bioinformatics/btaa579.
68. Björling E, Uhlén M. Antibodypedia, a portal for sharing antibody and antigen validation data. *Mol Cell Proteomics*. 2008;7:2028–2037. https://doi.org/10.1074/mcp.M800264-MCP200.
69. Hoffmann D, Haapasalo A. Patient-derived skin fibroblasts as a model to study frontotemporal lobar degeneration. *Neural Regen Res*. 2002;17:2669–2671. https://doi.org/10.4103/1673-5374.335814.
70. Morello G, La Cognata V, Guarnaccia M, La Bella V, Conforti FL, Cavallaro S. A diagnostic gene-expression signature in fibroblasts of amyotrophic lateral sclerosis. *Cells*. 2023;12:1884. https://doi.org/10.3390/cells12141884.
71. Alqahtani T et al. Mitochondrial dysfunction and oxidative stress in Alzheimer's disease, and Parkinson's disease, Huntington's disease and Amyotrophic Lateral Sclerosis—an updated review. *Mitochondrion*. 2023;71:83–92. https://doi.org/10.1016/j.mito.2023.05.007.
72. Olesen MA, Villavicencio-Tejo F, Quintanilla RA. The use of fibroblasts as a valuable strategy for studying mitochondrial impairment in neurological disorders. *Transl Neurodegener*. 2022;11:36. https://doi.org/10.1186/s40035-022-00308-v.
73. Katz DM et al. Preclinical research in Rett syndrome: setting the foundation for translational success. *Dis Model Mech*. 2012;5:733–745. https://doi.org/10.1242/dmm.011007.
74. de Mello CPP, Rumsey J, Slaughter V, Hickman JJ. A human-on-a-chip approach to tackling rare diseases. *Drug Discov Today*. 2019;24:2139–2151. https://doi.org/10.1016/j.drudis.2019.08.001.
75. Wu J, Yang Y, Yu J, Qiao L, Zuo W, Zhang B. Efficacy and safety of compassionate use for rare diseases: a scoping review from 1991 to 2022. *Orphanet J Rare Dis*. 2023;18:368. https://doi.org/10.1186/s13023-023-02978-x.
76. Benjin X, Ling L. Developments, applications, and prospects of cryo-electron microscopy. *Protein Sci*. 2020;29:872–882. https://doi.org/10.1002/pro.3805.
77. Terwilliger TC et al. AlphaFold predictions are valuable hypotheses and accelerate but do not replace experimental structure determination. *Nat Methods*. 2024;21:110–116. https://doi.org/10.1038/s41592-023-02087-4.
78. Tsirigotaki A et al. Mechanism of receptor assembly via the pleiotropic adipokine Leptin. *Nat Struct Mol Biol*. 2023;30:551–563. https://doi.org/10.1038/s41594-023-00941-9.
79. Park SH, Zhu P-P, Parker RL, Blackstone C. Hereditary spastic paraplegia proteins REEP1, spastin, and atlastin-1 coordinate microtubule interactions with the tubular ER network. *J Clin Invest*. 2010;120:1097–1110. https://doi.org/10.1172/JCI40979.
80. Sheppard CJR. The development of microscopy for super-resolution: confocal microscopy, and image scanning microscopy. *Appl Sci*. 2021;11:8981. https://doi.org/10.3390/app11198981.
81. Nixon-Abell J et al. Increased spatiotemporal resolution reveals highly dynamic dense tubular matrices in the peripheral ER. *Science*. 2016;354, aaf3928. https://doi.org/10.1126/science.aaf3928.
82. Saffari A et al. High-content small molecule screen identifies a novel compound that restores AP-4-dependent protein trafficking in neuronal models of AP-4-associated hereditary spastic paraplegia. *Nat Commun*. 2024;15:584. https://doi.org/10.1038/s41467-023-44264-1.
83. Khundadze M et al. A hereditary spastic paraplegia mouse model supports a role of ZFYVE26/SPASTIZIN for the endolysosomal system. *PLoS Genet*. 2013;9, e1003988. https://doi.org/10.1371/journal.pgen.1003988.
84. Accardi MV, Pugsley MK, Forster R, Troncy E, Huang H, Authier S. The emerging role of in vitro electrophysiological methods in CNS safety pharmacology. *J Pharmacol Toxicol Methods*. 2016;81:47–59. https://doi.org/10.1016/j.vascn.2016.03.008.
85. Liaci C et al. Loss of ARHGAP15 affects the directional control of migrating interneurons in the embryonic cortex and increases susceptibility to epilepsy. *Front Cell Dev Biol*. 2022;10, 875468. https://doi.org/10.3389/fcell.2022.875468.
86. Orrell RW. ALS2-related disorder [Internet]. *GeneReviews*. Seattle (WA): University of Washington, Seattle; 1993.
87. Hadano S et al. A gene encoding a putative GTPase regulator is mutated in familial amyotrophic lateral sclerosis 2. *Nat Genet*. 2001;29:166–173. https://doi.org/10.1038/ng1001-166.
88. Pardridge WM. Blood-brain barrier delivery. *Drug Discov Today*. 2007;12:54–61. https://doi.org/10.1016/j.drudis.2006.10.013.
89. Roy D, Hinge VK, Kovalenko A. To pass or not to pass: predicting the blood-brain barrier permeability with the 3D-RISM-KH molecular solvation theory. *ACS Omega*. 2019;4:16774–16780. https://doi.org/10.1021/acsomega.9b01512.
90. Abbott NJ, Dolman DEM, Patabendige AK. Assays to predict drug permeation across the blood-brain barrier, and distribution to brain. *Curr Drug Metab*. 2008;9:901–910. https://doi.org/10.2174/138920008786485182.
91. Helms HC et al. In vitro models of the blood-brain barrier: an overview of commonly used brain endothelial cell culture models and guidelines for their use. *J Cereb Blood Flow Metab*. 2015;36:862–890. https://doi.org/10.1177/0271678X16630991.
92. Urich E, Lazic SE, Molnos J, Wells I, Freskgård PO. Transcriptional profiling of human brain endothelial cells reveals key properties crucial for predictive in vitro blood-brain barrier models. *PLoS One*. 2012;7:e38149. https://doi.org/10.1371/journal.pone.0038149.
93. Siow S-F, Yeow D, Rudaks LI, Jia F, Wali G, Sue CM, Kumar KR. Outcome measures and biomarkers for clinical trials in hereditary spastic paraplegia: a scoping review. *Genes (Basel)*. 2023;14:1756. https://doi.org/10.3390/genes14091756.
94. Califf RM. Biomarker definitions and their applications. *Exp Biol Med*. 2018;243:213–221. https://doi.org/10.1177/1535370217750088.
95. Wilson DM, Cookson MR, Van Den Bosch L, Zetterberg H, Holtzman DM, Dewachter I. Hallmarks of neurodegenerative diseases. *Cell*. 2023;186:693–714. https://doi.org/10.1016/j.cell.2022.12.032.
96. Denton K et al. Impaired mitochondrial dynamics underlie axonal defects in hereditary spastic paraplegias. *Hum Mol Genet*. 2018;27:2517–2530. https://doi.org/10.1093/hmg/ddy156.
97. Kakouri AC et al. Transcriptomic characterization of tissues from patients and subsequent pathway analyses reveal biological pathways that are implicated in spastic ataxia. *Cell Biosci*. 2022;12:29. https://doi.org/10.1186/s13578-022-00754-1.

98. Åkesson J et al. Proteomics reveal biomarkers for diagnosis, disease activity and long-term disability outcomes in multiple sclerosis. *Nat Commun*. 2023;14:6903. https://doi.org/10.1038/s41467-023-42682-9.
99. Gautam M et al. 2-year-old and 3-year-old Italian ALS patients with novel ALS2 mutations: identification of key metabolites in their serum and plasma. *Metabolites*. 2022;12:174. https://doi.org/10.3390/metabo12020174.
100. Lesca G et al. Infantile ascending hereditary spastic paralysis (IAHSP): clinical features in 11 families. *Neurology*. 2003;60:674–682. https://doi.org/10.1212/01.wnl.0000048207.28790.25.
101. Eymard-Pierre E et al. Infantile-onset ascending hereditary spastic paralysis is associated with mutations in the alsin gene. *Am J Hum Genet*. 2002;71:518–527. https://doi.org/10.1086/342359.
102. Hadano S et al. Functional links between SQSTM1 and ALS2 in the pathogenesis of ALS: cumulative impact on the protection against mutant SOD1-mediated motor dysfunction in mice. *Hum Mol Genet*. 2016;25:3321–3340. https://doi.org/10.1093/hmg/ddw180.
103. Sato K et al. Altered oligomeric states in pathogenic ALS2 variants associated with juvenile motor neuron diseases cause loss of ALS2-mediated endosomal function. *J Biol Chem*. 2018;293:17135–17153. https://doi.org/10.1074/jbc.RA118.003849.
104. Hsu F, Spannl S, Ferguson C, Hyman AA, Parton RG, Zerial M. Rab5 and Alsin regulate stress-activated cytoprotective signaling on mitochondria. *eLife*. 2018;7:e32282. https://doi.org/10.7554/eLife.32282.
105. Boehm M, Aguilar RC, Bonifacino JS. Functional and physical interactions of the adaptor protein complex AP-4 with ADP-ribosylation factors (ARFs). *EMBO J*. 2001;20:6265–6276. https://doi.org/10.1093/emboj/20.22.6265.
106. Gadbery JE et al. Integrating structural and evolutionary data to interpret variation and pathogenicity in adapter protein complex 4. *Protein Sci*. 2020;29:1535–1549. https://doi.org/10.1002/pro.3870.