Fundamental Research 2 (2022) 851–858

Contents lists available at ScienceDirect

Fundamental Research

journal homepage: http://www.keaipublishing.com/en/journals/fundamental-research/

Review

Medical care of rare and undiagnosed diseases: Prospects and challenges

Zhiyan Shan[a,b], Lijun Ding[a,c], Caiyun Zhu[d], Ruijuan Sun[a,*], Wei Hong[a,*]

[a] Department of Health Sciences, National Natural Science Foundation of China, Beijing 100085, China
[b] Department of Histology and Embryology, Harbin Medical University, Harbin 150081, China
[c] Center for Reproductive Medicine and Obstetrics and Gynecology, Affiliated Drum Tower Hospital of Nanjing University Medical School, Nanjing, Jiangsu 210008, China
[d] State Key Laboratory of Medical Molecular Biology, Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences, School of Basic Medicine Peking Union Medical College, Beijing 100005, China

ARTICLE INFO

Article history:
Received 28 June 2022
Received in revised form 18 August 2022
Accepted 23 August 2022
Available online 7 September 2022

Keywords:
Rare disease
Undiagnosed disease
Gene
Animal model
Diagnosis
Therapy

ABSTRACT

Rare and undiagnosed diseases tend to be diverse, misdiagnosed, and difficult to diagnose. In some cases, the disease is progressive and life-threatening. Yet, to date, an estimated 95% of rare diseases have no approved therapy. Therefore, rare and undiagnosed diseases are considered the ultimate challenges for understanding human diseases. Here, we review the research progress, research frontiers, and important scientific issues related to rare and undiagnosed diseases. We mainly focus on five topics: (1) the identification and functional analysis of disease-causing genes; (2) the construction of cells, organoids, and animal models for mechanism validation; (3) subtyping and diagnosis; (4) treatment and drug screening based on causative genes and mutations; and (5) new technologies and methods for studying rare and undiagnosed diseases. In this review, we briefly update and discuss the pathogenic mechanisms and precision medicine for rare and undiagnosed diseases.

Rare and undiagnosed diseases are recognized as major public health concerns worldwide. It is estimated that over 250 million people worldwide, including 20 million individuals in China, suffer from rare and undiagnosed diseases. Misdiagnosis, undiagnosis, and a lack of appropriate therapy result in a huge emotional and financial burden on patients, their families, and the economy. An in-depth study of its pathogenesis will provide fundamental support for the accurate diagnosis and treatment of rare and undiagnosed diseases. This will contribute to the development of global medical science and new drugs.

1. Definition, challenges, and opportunities in the field of rare and undiagnosed diseases

Rare diseases are defined by the World Health Organization (WHO) as diseases with a prevalence of 0.65–1‰ [1]. In the United States, a rare disease is defined as a condition affecting fewer than 200,000 people, while in Europe, it is defined as a disease with an occurrence rate of less than 1/2000 [2]. The definition of rare diseases is important for not only solving medical problems but also providing a basis and scope for responding to public health problems and formulating policies. In the past few decades, China has strengthened its support for rare diseases. In 2010, the Medical Genetics Branch of the Chinese Medical Association, for the first time, defined rare diseases as any disease with a prevalence of less than 1/500,000 or a birth incidence of less than 1/10,000. To strengthen the management of rare diseases and safeguard the health rights and interests of patients, in 2018, the National Health Commission, Ministry of Science and Technology, Ministry of Industry and Information Technology, State Drug Administration, and State Administration of Traditional Chinese Medicine jointly published "China's First National List of Rare Diseases" with 121 rare diseases, most of which were undiagnosed. The released 1st national list promoted the rapid progression of China's rare disease diagnosis and treatment, basic research, drug development, and medical security. In 2019, China established a diagnosis and treatment network involving 324 hospitals through regional collaborations and remote support. This further improved the comprehensive diagnoses and treatment capacities for rare diseases [3]. In 2021, the definition of rare diseases in China was updated as a condition satisfying at least one of the following three criteria: an incidence of less than 1/10, 000 among newborns, a prevalence of less than 1/10,000, and an affected population of less than 140,000. Currently, there are approximately 7000 known rare diseases that together affect 10% of the national population. Approximately 80% of these

* Corresponding authors.
E-mail addresses: sunrj@nsfc.gov.cn (R. Sun), hongwei@nsfc.gov.cn (W. Hong).

https://doi.org/10.1016/j.fmre.2022.08.018

diseases have a genetic basis and are Mendelian disorders. Undiagnosed diseases are described as conditions eluding diagnosis, referring to phenotypically well-described diseases with unknown molecular bases, or complex phenotypic variations of common or rare diseases with heterogeneity, the mechanisms of which have not yet been elucidated. Notably, not all undiagnosed diseases are rare, indicating an intersection between the two categories.

Impressively, rare diseases occur frequently in young patients. According to the "China Rare Disease Industry Report" released by the Boston Consulting Group, approximately 60% of Chinese patients with rare diseases are under 20 years of age. Although some patients respond to treatment, only approximately 3% of patients can survive over 60 years of age. Furthermore, the shortage and poor accessibility of orphan drugs have been longstanding issues in China. Notably, only 5–10% of the approximately 7000 rare diseases are druggable. Of the 121 rare diseases on China's First National List of Rare Diseases, 47 lack treatment drugs worldwide, 16 have treatment drugs but are unavailable in China, and only 36 have approved pharmacotherapy in China. Considering the ethnic differences in genetics, environment, and other factors, there is an urgent need to develop effective therapies for rare diseases that can be applied to the Chinese population.

The released list of rare diseases has greatly accelerated the process of drug development and approval. International pharmaceutical companies have taken the initiative to accelerate the layout in the Chinese market. Notably, orphan drugs have been favored by many institutions because of their clear targets, low requirements for the number of enrolled cases, and high approval ratios. Currently, the development of orphan drugs is promoted, and some are generally available for application in common diseases (new indications). Meanwhile, multidisciplinary collaboration, multi-omics data analysis, the development of gene and cell therapy, and systematic parsing of disease phenotypes have provided unprecedented opportunities to improve the accurate diagnosis and treatment of rare and undiagnosed diseases.

2. Frontiers in rare and undiagnosed diseases

2.1. Identification and functional analysis of disease-causing genes for rare and undiagnosed diseases

The current dilemma in rare disease research lies in the rapid and effective identification of pathogenic mutations and the development of site-specific drugs. The etiology of rare and undiagnosed diseases is complicated, with a high degree of genetic and clinical heterogeneity. Generally, genetic variants are critical for rare and undiagnosed diseases. However, to date, only a small population of rare diseases have identified unusual genetic mutations, while the pathogenesis of most other rare diseases is still unknown. Additionally, in the absence of a comprehensive analysis of the connection between signaling pathways and phenotypes, the molecular mechanisms of rare diseases remain to be further validated. Therefore, efforts to identify novel causative mechanisms will markedly improve subtyping and innovation of therapeutic strategies.

In recent decades, several molecular strategies have been employed to uncover disease-related genes. First, genetic engineering and sequencing can help identify pathogenic variants. In addition, DNA polymorphic variation provides information on haplotype structure, linkage disequilibrium, hot spots for recombination, and population-specific mutational spectrum. For example, using a small set of polymorphic markers, the genetic variant for Huntington's disease was mapped to chromosome 4 in a large pedigree linkage analysis [4]. In addition, the development of genome infrastructure also promotes positional cloning precision. For instance, Xia et al. successfully cloned the *GJB3* gene using a homologous expressed sequence tag search and revealed that mutations in *GJB3* might be responsible for hereditary deafness [5]. Furthermore, the completion of the euchromatic fraction of the human genome has remarkably facilitated disease gene discovery. By combining linkage analysis and sequence identification, a collaborative team led by Liu and Zhang successfully discovered and cloned the causative gene *SLC20A2*, which is responsible for familial idiopathic basal ganglia calcification [6].

Previous studies on rare diseases have focused on single-base mutations or copy number variants (CNVs) in the coding region, whereas only a few studies have evaluated the influence of other types of *de novo* mutations, such as the repetition or deletion of large fragments. This is mainly limited by second-generation sequencing technology, which is ineffective for analyzing highly heterozygous genomes, highly repetitive sequences, high GC regions, complex CNVs, and large structural variations (SVs). Therefore, third-generation sequencing capable of producing substantially longer reads will help solve challenges with traditional genomic sequencing. Finally, it will help identify more causative genes and uncover the mechanisms in rare and undiagnosed diseases. In 2009, Ng et al. published the first report on whole exome sequencing (WES). They observed rare and common variants of *MYH3* and defined it as a disease-causing gene in patients with Freeman-Sheldon syndrome [7]. Combining WES with gene chip localization technology, Tang et al. team successfully identified the cloned SCA35 pathogenic gene *TGM6* [8]. Furthermore, the application of WES in a large family-based study (n = 4293) revealed that approximately 42% of patients with developmental disorders harbored *de novo* pathogenic mutations [9]. Over the past decade, WES has contributed to 25–45% of diagnoses, indicating its strong diagnostic potential for genetic disorders [10–13]. However, most of the pathogenic mechanisms require further validation. In 2021, a pilot study published in the New England Journal of Medicine involved analysis of the gene profiles of 4660 participants from the 100,000 Genomes Project and revealed that whole-genome sequencing (WGS) helps improve the efficiency of genomic diagnostics across a range of rare diseases. For instance, 13% rare diseases were caused by mutations in non-coding sequence or mitochondrial genomes, tandem repeat expansions in persons with Huntington's disease and others, while an additional 2% of the diagnoses involved in coding variants in regions of low coverage on exome sequencing. This pilot study supported that WGS is useful for discovering more genomic variants and that long-read sequencing may be more promising for rare disease diagnoses [14]. Fully phased long-read sequencing is more effective in detecting structural variations and acquiring sequence information that is poorly captured by short-read sequencing. Additionally, recent publications have highlighted the success of employing transcriptomic data provided by RNA-Seq to improve diagnostic yields in a wide range of rare diseases. By providing direct insights into the transcriptional alterations unique to patients, RNA-seq sheds light on the causative variants and thus, improves rare disease diagnosis beyond WES/WGS sequencing [15]. Therefore, it is valuable to explore the combination of traditional genomics techniques and RNA-Seq in both diagnostic and research settings. Hopefully, it will help to promote the overall diagnostic rates and provide new strategies for defining disease-causing mechanisms and diagnosis of rare diseases.

Although most rare diseases are characterized by Mendelian inheritance, pathogenic genes have only been identified in 20–40% of these conditions. However, non-coding sequences, make up 98% of the human genome and have been of functional importance in recent studies [16]. Non-coding regions include various regulatory elements and ncRNA genes, which affect gene expression, protein products, and metabolite levels depending on the regulation of transcription, translation, or chromatin homeostasis [17,18]. Recently, pathogenic variants located in the promoter, RNA-binding region, 5′ untranslated region (5′UTR), and non-coding SVs have been identified and verified to increase the risk of developing disorders, suggesting that non-coding pathogenic variants may be a novel interpretation of the loss of heredity [19–22]. In addition, epigenetic modifications, gene-environment interactions, and developmental compensatory mechanisms also affect inherited penetrance due to different changes in transcription, metabolism, or signaling pathways. For example, more than 68% of individuals with autism spec-

trum disorder (ASD) shared the same pattern of histone acetylation in the brain. This epigenetic pattern affecting common molecular pathways may also be responsible for the tremendous heterogeneity of ASD [23]. Similarly, using DNA methylation chips, Erfan et al. analyzed genomic methylation in patients with neurodevelopmental abnormalities and predicted epigenetic features that were unique to patients [24]. Genome-wide association studies (GWAS) have been used as a major tool for identifying genotype-phenotype associations in recent years. Most disease-associated genetic variants found using GWAS are located in non-coding regions of the genome [25]. Specifically, Gandal et al. reported a transcriptome-wide association study that integrated transcriptomic and genomic data from over 2000 brain samples from patients with schizophrenia, bipolar disorder, and ASD, and found that nearly 1000 ncRNAs are dysregulated in at least one disorder, and many have significant central nervous system enrichment without functional annotation [26]. Overall, from a systems biology perspective, the integrated analysis of multi-omics data, including genomics, transcriptomics, epigenetics, proteomics, and metabolomics, reflects the complexity of complicated diseases and provides new insights into the molecular etiology.

The lack of understanding of genetic patterns also limits etiology validation in some rare diseases. Besides monogenic disorders (classical Mendelian), polygenic and oligogenic disorders have attracted increasing attention. Oligogenic disorders are genetic diseases caused or modulated by multiple gene variants. For instance, Jing et al. suggested that oligogenic inheritance may play a role in the onset of pulmonary arterial hypertension (PAH) and that several genes, including *BMPR2, BMP9,* and *PTGIS*, may be responsible for PAH phenotypic variability. Hypermethylation of the *BMPR2* promoter, *MicroRNA-483,* and *DNMT3B*, were revealed to be involved in the epigenetic regulations in PAH [27–31]. Overall, it is recommended that more inheritance patterns warrant further attention by the in-depth discovery of new pathogenic genes, creating oligogenic inheritance maps of rare diseases, capturing key target molecules, and developing more effective treatment approaches.

Complex genetic mechanisms, such as non-Mendelian inheritance, non-coding variants, and epigenetic modifications, may be key factors in rare and undiagnosed diseases of unknown etiology. Currently, there are still several limitations to the identification and functional analysis of disease-causing genes. For example, it remains a big challenge to identify new pathogenic genes and mutations, understand inheritance patterns, reveal inherent and pathogenic mechanisms, and explore treatment strategies of genetic variants. In addition, deciphering rare and undiagnosed disease requires the development of new multidisciplinary technologies. The interdisciplinary collaboration of emerging technologies, such as third-generation sequencing, bioinformatics, structural biology, gene editing, and many other disciplines, has greatly deepened the understanding of pathogenic mechanisms and promoted the exploration of treatment strategies. Furthermore, the identification and functional analysis of genetic variants depend on the establishment of large-scale disease cohorts and large population reference databases. Therefore, the construction of national special disease cohorts (including family and high-quality cohorts) and biological sample banks may be valuable. In addition, national data-sharing networks are critical for the discovery of new pathogenic genes and mutants unique to patients with rare and undiagnosed diseases. Taken together, the key scientific issues associated with the management of rare and undiagnosed diseases are identifying new pathogenic genes and new variants, uncovering the pathogenic and molecular mechanisms, underlying genotype and phenotype associations in rare and undiagnosed diseases.

2.2. Construction and mechanism of cells, organoids, and animal models

With the development of sequencing technology and advancements in research methods, many new pathogenic genes and mutation loci have been identified. The gold standard for identifying pathogenic genes or loci is to accurately model disease phenotypes in genetic animal models. However, these models often do not faithfully replicate human diseases. Therefore, there is an urgent need to establish viable disease models for solving clinical issues in rare disease research, such as high population prevalence, the low prevalence in single diseases, scarcity of clinical samples, difficulty in cohort establishment, challenges in performing clinical trials, and sluggish traditional screening. Therefore, the establishment of stable and reliable disease models is indispensable for studying pathogenic mechanisms and developing therapies.

2.2.1. Cell and organoid models of rare diseases

Stem cells have the potential to differentiate into multiple functional cell types, paving the way for the treatment of genetic diseases. Induced pluripotent stem cells (iPSCs) are currently one of the most promising tools. iPSCs carrying the patients' genetic information can be stored for a long time and are convenient for the establishment and accumulation of clinical cohorts. iPSCs may help to solve the problem of spatiotemporal limitations in research on rare diseases. Currently, cells, tissues, and organoids derived from patient iPSCs are important models for studying rare and undiagnosed diseases. For instance, a study of xeroderma pigmentosum (XP), a rare hereditary skin condition, showed that neural cells derived from XP patient-specific iPSCs were hypersensitive to DNA damage-induced apoptosis [32]. This suggests that XP-iPSCs are valuable tools for clarifying the molecular mechanisms underlying neurological abnormalities in patients with XP. Similarly, Liu et al. reported that vascular smooth muscle cells induced from progeria children-specific iPSCs manifested aging-related phenotypes, including nuclear contraction, heterochromatin deletion, and telomere depletion [33]. As approximately 80% of rare diseases are caused by a single-base mutation, single-base editing technology has become a popular tool for gene therapy in rare diseases, such as spinal muscular dystrophy, thalassemia, and hemophilia. Since iPSCs can be easily generated from gene-edited single colonies, iPSCs and their organoids are emerging as one of the most important tools for novel cell or gene therapy. Xie et al. reported that the *HBB* gene corrected in β-thalassemia iPSCs using the CRISPR/Cas9 tool and revealed that compared with parental iPSCs, gene-corrected iPSCs restored the expression of *HBB* [34]. Additionally, Chinese researchers have assessed the roles of engineered iPSC-derived regulatory immune cells, specifically macrophages, and proposed engineered macrophages for use in treating complicated autoimmune diseases such as systemic lupus erythematosus. Patient-derived organoids have unique advantages for rapid disease modeling and drug screening in rare disease studies. Amyotrophic lateral sclerosis (ALS) is a rare disease with a high incidence rate. Traditional research approaches mostly focus on typical pathogenic genes (such as *TDP43*), and it is difficult to find the comorbid mechanism of ALS patients, which results in a long disease course. Recently, Pereira et al. established human ALS-iPSCs and derived sensorimotor organoids that contained physiologically functional neuromuscular junctions. They demonstrated that multidimensional information, including model building, pathogenic diagnosis, neurophysiological detection, and multi-omics screening, can be completely collected in one year [35]. Their work greatly accelerated organoid-related research and helped explore the comorbid mechanisms of ALS. Samarasinghe et al. generated organoids from iPSCs from patients with Rett syndrome and revealed abnormal epileptiform network activity in organoids, which was rescued with the unconventional neuromodulator drug pifithrin-α. This revealed the potential value of characterizing such rare conditions and identifying individualized therapeutic drugs [36]. iPSCs enable the development of an unlimited cell source for any human tissue and are feasible for fast and high-throughput expansion. Recent studies have reported the successful generation of iPSCs or organoids by reprogramming patient-specific cell sources, such as blood, skin, and even urine. Combining advanced technologies, including gene editing, cell factories, and high-throughput screening, will greatly accelerate drug discovery by constructing an iPSC-based high-throughput screening system.

2.2.2. Animal models

Animal models play a pivotal role in disease research. This is indispensable for identifying causative genes, revealing pathogenic mechanisms, identifying drug targets, and screening and validating therapeutic drugs. Moreover, the availability of suitable animal models is vital for evaluating therapeutic effects. An ideal animal model is characterized by accurate modeling human disease, predictability of drug treatment, ease of experimental manipulation, and low costs.

Currently available animal models are mostly rodents. Various transgenic animal models with genetic mutations have been established to mimic rare human diseases. For example, Fgf23-R176Q transgenic mice effectively mimic autosomal dominant hypophosphatemic rickets accompanied by hypophosphatemic osteomalacia when fed a low-iron diet [37]. Glycogen storage disease type Ia is caused by a lack of glucose-6-phosphatase-α (G6pc) activity. The G6pc knockout mouse model showed severely disturbed glucose homeostasis and rarely lived beyond weaning, which is similar to the pathological characteristics of humans [38]. Notably, the National Resource Center of Model Mice plans to provide approximately 100 mouse models for rare disease research at cost prices in the following years, which would greatly promote the development of rare disease research. Furthermore, because of the significant evolutionary differences between rodents and humans, non-human primates and trees show better prospects for mimicking reality for designing animal models. Chen et al. reported that *MeCP2* mutant monkeys mimic the clinical phenotype, advanced cognition, and disease process of ASD [39,40]. As mentioned above, monkey models provide effective study tools for intervention strategies, including elucidating molecular mechanisms and exploring gene therapy. In a recent non-human primate study, the investigators used adenine base editing to knock-down *PCSK9* in cynomolgus monkeys and demonstrated 66% whole-liver *PCSK9* editing and an approximately 60% reduction in blood LDL-C levels [41]. This study provides preclinical evidence of precise *in vivo* editing for non-human primates and provides novel insights into the prevention of rare diseases.

2.2.3. Semi-cloning technology

Li et al. reported the creation of androgenetic haploid embryonic stem cells (AG-haESCs) that could support full-term embryonic development upon injection into MII oocytes, leading to the generation of semi-cloned (SC) mice [42]. Importantly, they further demonstrated that by combining the CRISPR-Cas9 library and AG-haESCs, SC mice carrying different mutant genes could efficiently be generated, enabling rapid phenotyping of uniform founders without mosaicism in one generation. AG-haESC-mediated semi-cloning technology provides a rapid and efficient experimental assay for identifying pathogenic genetic variants. For instance, Müllerian anomalies (MA) are considered complex congenital diseases because no major genes have been found to account for human MA in monogenic inheritance. Through the joint utilization of genetic analysis and semi-cloning technology, researchers confirmed that single genetic variants in *GEN1* or *WNT9B* were insufficient for MA, whereas double heterozygous mutations in *GEN1* and *WNT9B* could lead to abnormal development of Müllerian ducts as a result of the synergistic mutational effect [43]. Numerous studies have shown that myotonic dystrophy type 1 (DM1) is a genetic disorder that compromises multiple organs, however, there is no suitable mouse model for mechanistic and potential drug screening studies. Li et al. successfully generated quadruple heterozygous mutant mice in one step by injecting haploid ESCs carrying quadruple mutant genes (*Dmpk, Six5, Mbnl1*, and *Dmwd*) into the oocytes. Mice with quadruple mutations can mimic symptoms of the most severe form of DM1 [44]. Prof. Richard H. Finnell also highlighted this work and appreciated that it represented an excellent proof of principle to support the combination of CRISPR-Cas9 and AG-haESCs for studying complex diseases [45]. Semi-cloning technology simplifies the recapitulation of complex human diseases, breaks the bottleneck of traditional models, and provides support for identifying pathogenic gene mutations and their combinations. This approach is vital for dissecting the underlying mechanisms of rare diseases and also provides new models and therapeutic strategies for exploring genetic etiology and potential treatment.

Taken together, based on construction and mechanistic research on cells, organoids, and animal models, the key scientific issues are systematically establishing disease-specific iPSCs and highly effective animal models. This will help recapitulate rare and undiagnosed human diseases, expand the production of gene-edited cell and animal models, and provide an important research platform for screening disease-causing genes, pathogenic mechanisms, and drug development.

2.3. Subtyping and diagnosis of rare and undiagnosed diseases

Rare diseases usually exhibit high phenotypic heterogeneity and unobvious syndromes at an early stage; therefore, they inevitably hinder both diagnosis and treatment. It has been reported that patients with rare diseases visit an average of eight physicians before receiving an accurate diagnosis. The average time to obtain a diagnosis of a rare disease is 5–7 years, and 2–3 incorrect diagnoses were made for a given condition [16]. A delayed diagnosis may lead to irreversible progression of the patient's condition. Pathological changes are the basis of current clinical diagnoses, which do not apply to undiagnosed diseases. Thus, clinical diagnosis patterns should shift from a classic phenotype-driven model to a genotype-driven one. The development of next-generation sequencing technology has made it possible to establish precise clinical diagnostic patterns. These measures may remodel the hierarchy, subtyping, and diagnosis patterns and promote the development of precision medicine for rare and undiagnosed diseases.

Recently, Yang et al. observed that inherited skin disorders are caused by defective ion channel functioning. They revealed that mutations in the sodium channel α subunit-coding *SCN9A* gene cause primary erythermalgia [46]. Similarly, gain-of-function mutations in *TRPV3*, a gene coding transient receptor potential ion channel, causes Olmsted syndrome characterized by palmoplantar and periorificial keratoderma, itch, and hair loss [47,48]. According to these studies, they proposed a new term "skin channelopathies" and developed drugs targeting ion channels. Undiagnosed diseases may be serious or life-threatening during childhood, and an early diagnosis for these children will provide direction for personalized treatments and improve their quality of life. For instance, subtyping based on genotype-phenotype combinations can anchor causative genes using only a few hundred samples from ASD patients with high heterogeneity. This greatly improves the efficiency of the identification of pathogenic genes.

The emphasis of rare disease subtyping and diagnosis is to establish genotype-phenotype correlations. Currently, some critical issues, such as identifying pathogenic genes and variants based on multi-omics data and determining molecular targets suitable for clinical diagnoses, need to be further explored. Single-omics techniques also help predict the symptoms of rare diseases. For instance, using blood transcriptome sequencing, Frésard et al. established a core network of rare disease-related genes and demonstrated how transcriptome data from a large control cohort. This can be directly utilized for rare disease identification, even prior to generating disease-specific differentiated cell types, further aiding the interpretation of risk variants in a unique cell type relevant for rare disease research [49]. In comparison, the integration of multi-omics data is more useful for diagnosing and revealing new pathogenic mechanisms of rare and undiagnosed diseases. For instance, Cummings et al. reported that the integration of transcriptome sequencing, WES, and WGS data from patients with undiagnosed muscle disorders yielded a higher diagnosis rate, reinforcing the power of a large systematic application to validate candidate splice-disrupting mutations [50]. Additionally, Snyder et al. integrated autism-related genomic, transcriptomic, and proteomic data and constructed a multi-layer module enriched for known genes, as well as genes harboring rare mutations in autism [51]. Furthermore, mapping rare variants in patients onto pro-

tein complexes revealed novel proteins and the molecular machinery involved in autism.

Currently, most existing large-scale studies are limited to European and American enrollment; therefore, it is necessary to establish national rare disease cohorts and disease-oriented biobanks in China and develop artificial intelligence strategies based on both genotypes and phenotypes. These innovations are essential for understanding disease mechanisms, establishing clinical diagnostic patterns based on a combination of phenotypes and genotypes, and developing personalized therapies.

Taken together, the key of subtyping and diagnosis is pinpointing complex molecular mechanisms by utilizing multi-omics datasets, remolding phenotype-genotype-driven diagnosis patterns, and developing personalized and accurate diagnoses of rare and undiagnosed diseases.

2.4. Treatment and drug screening based on causative genes and mutations

The pathogenesis of rare and undiagnosed diseases is complex. Due to the limited understanding of most rare diseases, the available treatments are not sufficient to meet the demands of the patients. Currently, new emerging novel therapies, such as small-molecule drugs, protein-based therapeutic agents (proteins, peptides, and antibodies), antisense oligonucleotides, small interfering RNA (siRNA), and cell and gene therapies, have driven the development of rare and undiagnosed diseases. To meet the demand for the clinical diagnosis of rare diseases, it is urgent to explore new therapeutic targets, develop new therapeutic drugs, and optimize gene therapy technologies.

2.4.1. Small-molecule compounds

The demand for cost-effective tissue-targeting therapies keeps small-molecule drugs at the forefront of discovery efforts and tends to be stable and easy to scale. Loss- or gain-of-function mutations in protein products are considered the root cause of rare diseases with well-defined pathogenic genes. Traditional drug discovery focuses on altering protein levels by direct inhibitors or activators, and hence, regulates disease states. Therefore, although thousands of causative proteins are involved in the development of rare diseases, 70–80% are categorized as "undruggable". Therefore, targeted intervention strategies are not applicable in most rare diseases. One appealing idea to address is to establish small molecule-targeted protein degradation technologies that may provide chemical biology tools for studying pathogenic mechanisms and developing therapeutic drugs and potentially revolutionizing biomedicine. To overcome these problems, Lu et al. developed innovative targeted therapeutic technologies based on small-molecule compounds, including autophagosome-tethering compounds to degrade pathogenic proteins and enhance the global protein quality control system. Moreover, they developed a new high-throughput compound screening platform based on small-molecule microarray [52–54]. These studies provided a new fundamental idea of intervention treatment, which is expected to establish a new paradigm for drug development in rare diseases.

2.4.2. Gene therapy

Gene therapies seek to correct DNA alterations responsible for any given genetic disease. Gene therapy is emerging as an effective treatment for genetic diseases, either by replacing defective genes or by inserting normal copies into the genome. In recent years, several gene therapy drugs have shown great potential in clinical settings [55]. For instance, SPK-9001 is an investigational bioengineered vector that expresses factor IX. Patients with hemophilia B in a phase 1/2 clinical trial of SPK-9001 experienced increased coagulant activity and reduced bleeding episodes [56]. Furthermore, Luxturna is a gene therapy drug for leber congenital amaurosis caused by mutations in *RPE65* [57]. It has been reported that one dose of luxturna could slow the progression of vision loss and restore some vision, particularly, night vision. Additionally, Onpattro® is a liposomal siRNA specifically targeting transthyretin (TTR) [58], which reduces the accumulation of TTR in tissues and results in improved treatment of polyneuropathy in patients with hereditary transthyretin amyloidosis. In recent years, gene therapy for rare diseases has been developed. Notably, the number of clinical trials focusing on eye disease alone has increased by more than 200%. In 2019, over 250 companies worldwide were engaged in gene therapy and gene-modified cell therapy, and nearly a thousand clinical projects are underway. However, gene therapy-related problems, such as safety risks, long-term effects, and the low efficiency of gene delivery, remain to be solved. Furthermore, some technological bottlenecks still limit gene therapy, such as precisely delivering exogenous genes to target cells, reducing tumor risk by insertional mutagenesis of integrated vectors, and ensuring long-term effectiveness as non-integrating vectors.

Notably, the efficacy of gene therapy in complex diseases is still not ideal. This is mainly due to the lack of identification of the core causative genes in rare diseases. Thus, it would be worthwhile to switch the direction of gene therapy by exploring the relationship between controlling the symptoms of the disease and the regulation network of causative genes, and not just focusing on the expression levels of these genes.

2.4.3. Novel gene therapy strategies based on gene editing

In recent years, gene editing technology has been used to precisely edit or correct genomes. Such technology is also expected to solve the shortcomings of the available gene therapies for rare diseases. CRISPR/Cas9, a new gene-editing tool, provides an important strategy for the precise treatment of rare diseases [59,60]. Currently, utilizing and innovating gene editing tools is critical for the development of gene therapy.

Newby et al. used a custom adenine base editor to convert the sickle cell disease (SCD) allele (HBB^S) into non-pathogenic β-globin (HBB^G). They observed a one-time autologous treatment for SCD in mice that eliminated pathogenic HBB^S and generated benign HBB^G [61]. In preclinical research, Liang et al. developed a novel gene therapy for hemophilia A (HA), including *in situ* genetic correction or non-viral vector delivery systems [62]. Impressively, FVIII secretion was rescued in candidate cell types derived from gene-corrected iPSCs. Additionally, they further explored gene-edited allogeneic stem cell therapy and CRISPR/Cas9 *in vitro* therapeutic strategies to promote the clinical development of gene therapy for HA. Recently, CRISPR Therapeutics and researchers at Stanford University have initiated clinical trials on sickle cell disease using different gene therapies. CRISPR Therapeutics used CRISPR turn on a fetal form of hemoglobin in patient's blood stem cells and reinfused edited blood stem cells into patients to correct sickle cell disease [63]. Porteus et al. outlined a CRISPR-based method for targeting hematopoietic stem cells via homologous recombination at the *HBB* locus to advance the development of next-generation therapies for β-hemoglobinopathies [64]. CRISPR gene editing-based therapies have also been approved for the investigation of human diseases such as melanoma, synovial sarcoma, and multiple myeloma.

Despite continuous improvements, gene editing technology still has some drawbacks, including off-target mutations, safety risks, a wider editing window of single-base editors, and the low editing efficiency of targeted integration of large fragments. To overcome these limitations, Ping et al. reported that the NanoProCas9 system integrates targeted delivery and the conditional activation of CRISPR/Cas9, avoiding off-target mutations at non-targeted sites and offering precise genome editing for inflammatory diseases [65,66]. Furthermore, Li et al. reported the identification and engineering of Cas12b, which is small and highly specific, making it suitable for therapeutic genome editing [67,68]. However, the safety and efficiency of this gene-editing system are challenging. Currently, the adeno-associated virus (AAV) is the most popular vector for *in vivo* applications [69]. The efficacy of *in vivo* CRISPR/Cas9-mediated substrate reduction therapies to treat primary hyperoxaluria type I (PH1) was evaluated; the administration of an AAV8-CRISPR/Cas9 vector prevented oxalate overproduction and kidney damage, with no signs of toxicity in $Agxt1^{-/-}$ mice [70]. The key issues of virus vector delivery include the development of more AAV serotypes, safer and more precise delivery, and whole-novel virus vec-

tor tools. Owing to their controllable physicochemical properties, simple preparation, and viable biocompatibility, non-viral delivery materials have remarkable potential in preclinical research and clinical trials. The rapid advancement of gene editing tools and gene delivery systems will accelerate the development of new strategies for gene therapy.

Taken together, there are several key scientific issues limiting the diagnosis and treat of rare diseases and must be solved through the following efforts: (1) identifying new therapeutic targets and accelerating the clinical translation of new pharmaceuticals, (2) innovative paradigms for new drug development and exploring new treatment options, (3) launching new trials to develop new patterns and strategies for gene therapy, improving the safety and effectiveness of gene therapy, and (4) modifying existing weak patterns or using drug repurposing to accelerate the qualitative development of the therapeutic field.

2.5. New technologies and methods for studying rare and undiagnosed diseases

The key aspect of the prevention of rare diseases is the identification of causative genes to facilitate early diagnosis and intervention. Recent progress in undiagnosed diseases has demonstrated that traditional options that focus on monogenic or multigenic research have not met the need for diagnosis and therapy. Therefore, interdisciplinary studies will provide more concepts and research patterns, help to elucidate pathogenic mechanisms and promote early warning and prevention of the pre-transition state of rare diseases.

Biomarkers play a vital role in disease detection, treatment, and follow-up. However, the expression of traditional biomarkers is not always constant with changes in time and conditions. In contrast, molecular networks represent a steady biological state. Network biomarkers comprise a group of molecules that are strongly correlated with each other. Therefore, network biomarkers as biological features have great advantages in terms of stability and accuracy. In 2018, Chen et al. developed a model-free method that theoretically derived an index based on a dynamical network biomarker (DNB). The application of a single DNB to clinical data shows that DNB members are effective for early warning and prognostic analysis [71]. Specifically, the DNB method successfully identified critical transitions in liver cancer and colorectal cancer based on genomics or proteomics data [72,73]. Therefore, DNB can predict critical transitions by identifying fluctuating features, quantitatively detecting the pre-transition state, serving as an effective early warning signal, and providing viable theory and methods for warning [74,75]. Moreover, before the critical transition occurs, early intervention or enhancement of the molecular functions of DNB may significantly delay sudden deterioration. This may be a potential entry point for breaking through the traditional idea of investigating complex disorders.

Various biomolecule modifications are also important factors that affect the occurrence and development of diseases. Glycosylation is an enzymatic process that transfers carbohydrates or glycans to proteins or other organic molecules, such as lipids. This modification is highly complex because glycosylation sites can be decorated with various sugar chains. Moreover, the lack of updated and convenient databases prevents storage and references to emerging glycosylation data. Alterations in site-specific glycosylation patterns have been associated with physiological functions. Therefore, aberrant glycosylation has been implicated as a significant contributor to many human diseases [76].

PANX1, a member of the pannexin family, is a highly glycosylated channel-forming protein. Wang et al. described four independent families in which mutations in *PANX1* caused familial or sporadic female infertility via oocyte death phenotype [77]. These mutations alter the PANX1 glycosylation pattern and result in aberrant PANX1 channel activity and ATP release in oocytes. This study showed the critical role of glycosylation patterns in human oocyte development, providing a genetic explanation for a subtype of human reproductive diseases. Glycans play essential functional roles in various biological processes and are one of the major types of biomacromolecules in cells. Due to the staggering complexity of glycans, the detection of glycan glycosylation is extraordinarily difficult. Therefore, strategies for labeling, imaging, and profiling glycans have been key issues in glycobiology research. Furthermore, Chen et al. developed precise chemical tools and intact glycopeptide analysis methods to probe glycosylation in a cell-specific and protein-specific manner [78,79]. The application of these methods enables visualization and tracking of glycans in cells, tissues, and living organisms. Therefore, dissecting glycosylation regulation is important for uncovering the roles of glycosylation in physiological processes, such as neural networks, cell heterogeneity, and embryonic development, and for improving disease diagnosis and therapeutics.

Taken together, the development of new technologies and methods is critical for investigating rare and undiagnosed diseases. Interdisciplinary systems will further accelerate the establishment of new technologies and innovative research paradigms, and elucidate the pathogenic mechanisms, ultimately improving the precise diagnosis and treatment of corresponding diseases.

3. Challenges and prospects of rare and undiagnosed diseases

3.1. Identification and functional analysis of pathogenic genes

Currently, the identification of new pathogenic genes and mutations in rare and undiagnosed diseases has much room for improvement. Therefore, for uncovering the pathogenic mechanisms and elucidating molecular mechanisms based on genotype-phenotype correlation, the following steps should be taken: (1) Systematic identification of new disease-causing genes and variants and revealing new genetic inheritance patterns. (2) Comprehensive exploration of the relationship between disease genotype and phenotype. (3) Interpretation of the pathogenic mechanisms based on the discovery of new pathogenic genes and variants. (4) Researchers should focus on common problems, such as studies of common mechanisms between different rare disorders, common diseases, and rare diseases.

3.2. Disease models of rare and undiagnosed diseases

Animal models and scarce non-human primates still limit the study of pathogenic mechanisms and the development of drugs for rare and undiagnosed diseases. The following steps should be taken: (1) Utilization of new technologies to achieve high-efficiency and large-scale development of effective animal models for understanding the pathogenesis of and discovering drug targets. (2) Systematic and standardized establishment of iPSCs and iPSC-derived cells and organoids for providing an effective platform for research on unveiling pathogenic mechanisms and drug screening. (3) Expanding the range of non-human primates to select suitable animal models for rare disease research, in addition to establishing primate models using single-base editing technology based on monogenic diseases.

3.3. Subtyping and diagnosis of rare and undiagnosed diseases

Early diagnosis and intervention and improved diagnostic rates of rare and undiagnosed diseases play an important role in promoting the clinical revolution of genomic medicine. New diagnostic patterns should be established based on the following: (1) Developing new diagnostic methods for undiagnosed diseases, integrating second- and third-generation sequencing technology, multi-omics technology, and remodeling disease diagnosis systems. (2) Clarifying disease subtyping by shifting the diagnosis patterns from classic phenotype-driven models to genotype-driven sub-molecular models. (3) Establishing phenotype-genotype-driven precise diagnosis patterns and developing novel research paradigms by building a large-scale population genotype database (including healthy and diseased populations), achieving accurate subtyping through a single-phenotype genome project, and achieving precise phenotypic tracking by building a multiyear follow-up co-

hort. (4) Comprehensive and systematic understanding of disease occurrence and development by dissecting spatial-temporal development in the same disease.

3.4. Treatment and drug screening of rare and undiagnosed diseases

To date, the following key issues limit the treatment and drug screening for rare disease therapy: most causative proteins are “undruggable”, delivery systems are less safe and efficient, and the long-term efficacy of drugs is difficult to maintain.

In the future, the following steps should be taken: (1) Uncovering new therapeutic targets based on new pathogenic genes or novel pathogenic mutations of known genes in rare diseases. (2) Improving gene editing technology, precise regulation technology for genome drugs and vector delivery systems of gene therapy, enhancing the targeting of organs/tissues and sub-organelles, optimizing the efficiency of gene therapy, making full use of the established non-human primate models based on single-gene diseases to evaluate the safety and efficacy of gene therapy. (3) Establishing innovative technology platforms and developing original gene therapy, cell therapy technologies, and targeted regulatory technologies.

3.5. New technology and method for studying rare and undiagnosed diseases

To fully integrate the resources of basic research and clinical platforms, collect and share standardized high-quality clinical and basic research data, and establish a collaborative group including basic and clinical research, the following steps should be taken: (1) Development of new technologies and methods to study the genetic factors involved in the occurrence and development of rare and undiagnosed diseases. (2) Application of DNB to reveal the mechanisms underlying the occurrence and development of disease to achieve early diagnosis and early signs of rare diseases. (3) Focus on biomolecular-specific modifications (such as glycosylation) and provide a theoretical basis for understanding disease pathogenesis, diagnosis, and treatment.

4. Concluding remarks and prospects

Multidisciplinary and interdisciplinary collaboration is vital for research on rare and undiagnosed diseases. The national foundation will encourage researchers to (1) take advantage of rich clinical resources and discover new pathogenic genes and variants of rare and undiagnosed diseases, (2) elucidate genotypes and phenotypes by applying suitable cell and animal models, (3) screen new targets for diagnosis and treatment, and establish diagnosis and treatment patterns.

To promote original research breakthroughs, the construction of an innovative ecology for clinical transformation, and the sustainable development of rare and undiagnosed diseases should be promoted. We propose that the following critical issues should be achieved: (1) Identification of new disease-causing genes and new variants and dissection their functional and molecular mechanisms. (2) Dissection of the genotype-phenotype relationship of rare and undiagnosed diseases, and elucidation of the molecular mechanism of the heterogeneity of complex undiagnosed diseases. (3) Establishment of animal models that efficiently and specifically mimic rare human diseases using multidisciplinary technologies. Systematic and standardized establishment of iPSC models for rare and undiagnosed diseases to provide effective platforms for unveiling pathogenic mechanisms and drug screening studies. (4) Uncover network biomarkers of rare and diagnosed diseases, develop accurate diagnosis and molecular subtyping patterns, and disease early warning systems based on genotype-phenotype correlations. (5) Discover new therapeutic targets based on causative genes and variants, mechanistically analyze rare and undiagnosed diseases, apply drug repurposing to screen potential drug candidates, and establish a precise diagnosis and treatment system based on phenotype-genotype integration. (6) Develop highly efficient gene therapy systems, in particular, precise modification techniques based on DNA- or RNA-editing technologies, effectively use endogenous vectors to optimize delivery systems, and improve the targeted delivery, effectiveness, and safety of gene therapy. (7) Strengthen the cooperation of multidisciplinary and multifield research, including clinical medicine, genetics, bioinformatics, genetic engineering, and chemical biology, and develop innovative research paradigms to promote the integration of new technologies and methods, including epigenetic factor identification, molecular network analysis, accurate diagnosis of large-scale disease data, glycosylation modification detection, and gene editing applications. (8) Build large-scale population genotype databases (including healthy and diseased populations) and multiyear follow-up cohorts to achieve resource sharing.

Author contributions

Z.Y.S. and C.Y.Z. conceived of the review. Z.Y.S. prepared drafts. W.H., R.J.S. and L.J.D. revised the manuscript. All authors contributed to the discussion. All authors have read and approved the final version of the manuscript.

Declaration of competing interest

The authors declare that they have no conflicts of interest in this work.

References

[1] P. Song, J. Gao, Y. Inagaki, et al., Rare diseases, orphan drugs, and their regulation in Asia: current status and future perspectives, Intractable Rare Dis. Res. 1 (2012) 3–9.
[2] J.J. Gagne, L. Thompson, K. O’Keefe, et al., Innovative research methods for studying treatments for rare diseases: methodological review, BMJ 349 (2014) g6802.
[3] China to establish collaboration network for rare disease diagnosis, treatment, 2019. http://english.www.gov.cn/state_council/ministries/2019/02/18/content_281476527383692.htm, 2019 (accessed 18 Feb 2019).
[4] J.F. Gusella, N.S. Wexler, P.M. Conneally, et al., A polymorphic DNA marker genetically linked to Huntington’s disease, Nature 306 (1983) 234–238.
[5] J.H. Xia, C.Y. Liu, B.S. Tang, et al., Mutations in the gene encoding gap junction protein beta-3 associated with autosomal dominant hearing impairment, Nat. Genet. 20 (1998) 370–373.
[6] C. Wang, Y. Li, L. Shi, et al., Mutations in SLC20A2 link familial idiopathic basal ganglia calcification with phosphate homeostasis, Nat. Genet. 44 (2012) 254–256.
[7] S.B. Ng, E.H. Turner, P.D. Robertson, et al., Targeted capture and massively parallel sequencing of 12 human exomes, Nature 461 (2009) 272–276.
[8] J.L. Wang, X. Yang, K. Xia, et al., TGM6 identified as a novel causative gene of spinocerebellar ataxias using exome sequencing, Brain 133 (2010) 3510–3518.
[9] Deciphering Developmental Disorders S, Prevalence and architecture of de novo mutations in developmental disorders, Nature 542 (2017) 433–438.
[10] Y. Yang, D.M. Muzny, J.G. Reid, et al., Clinical whole-exome sequencing for the diagnosis of mendelian disorders, N. Engl. J. Med. 369 (2013) 1502–1511.
[11] Y. Yang, D.M. Muzny, F. Xia, et al., Molecular findings among patients referred for clinical whole-exome sequencing, JAMA 312 (2014) 1870–1879.
[12] H. Lee, J.L. Deignan, N. Dorrani, et al., Clinical exome sequencing for genetic identification of rare Mendelian disorders, JAMA 312 (2014) 1880–1887.
[13] S.E. Soden, C.J. Saunders, L.K. Willig, et al., Effectiveness of exome and genome sequencing guided by acuity of illness for diagnosis of neurodevelopmental disorders, Sci. Transl. Med. 6 (2014) 265ra168.
[14] G.P.P. Investigators, D. Smedley, K.R. Smith, et al., 100,000 Genomes pilot on rare-disease diagnosis in health care - preliminary report, N. Engl. J. Med. 385 (2021) 1868–1880.
[15] Monica Heger, ASHG: RNA Sequencing can improve diagnostic yield of exome, genome tests. From Genome Web. https://www.rna-seqblog.com/ashg-rna-sequencing-can-improve-diagnostic-yield-of-exome-genome-tests/, 2018 (accessed 23 Oct 2018).
[16] R.E. Pogue, D.P. Cavalcanti, S. Shanker, et al., Rare genetic diseases: update on diagnosis, treatment and online resources, Drug Discov. Today 23 (2018) 187–195.
[17] L.I. Patrushev, T.F. Kovalenko, Functions of noncoding sequences in mammalian genomes, Biochemistry (Mosc) 79 (2014) 1442–1469.
[18] O.M. de Goede, D.C. Nachun, N.M. Ferraro, et al., Population-scale tissue transcriptomics maps long non-coding RNAs to complex disease, Cell 184 (2021) 2633–2648 e19.
[19] J.M. Havrilla, B.S. Pedersen, R.M. Layer, et al., A map of constrained coding regions in the human genome, Nat. Genet. 51 (2019) 88–95.
[20] D. Vitsios, R.S. Dhindsa, L. Middleton, et al., Prioritizing non-coding regions based on human genomic constraint and sequence context with deep learning, Nat. Commun. 12 (2021) 1504.

[21] J. Zhou, C.Y. Park, C.L. Theesfeld, et al., Whole-genome deep-learning analysis identifies contribution of noncoding mutations to autism risk, Nat. Genet. 51 (2019) 973–980.
[22] E. Turro, W.J. Astle, K. Megy, et al., Whole-genome sequencing of patients with rare diseases in a national health system, Nature 583 (2020) 96–102.
[23] W. Sun, J. Poschmann, R. Cruz-Herrera Del Rosario, et al., Histone acetylome-wide association study of autism spectrum disorder, Cell 167 (2016) 1385–1397 e11.
[24] E. Aref-Eshghi, E.G. Bend, S. Colaiacovo, et al., Diagnostic Utility of Genome-wide DNA Methylation Testing in Genetically Unsolved Individuals with Suspected Hereditary Conditions, Am. J. Hum. Genet. 104 (2019) 685–700.
[25] E. Rojano, J.A. Ranea, J.R. Perkins, Characterisation of non-coding genetic variation in histamine receptors using AnNCR-SNP, Amino Acids 48 (2016) 2433–2442.
[26] M.J. Gandal, P. Zhang, E. Hadjimichael, et al., Transcriptome-wide isoform-level dysregulation in ASD, schizophrenia, and bipolar disorder, Science 362 (2018) eaat8127.
[27] J. Zhang, Y. He, X. Yan, et al., MicroRNA-483 amelioration of experimental pulmonary hypertension, EMBO Mol. Med. 12 (2020) e11303.
[28] Y. Yan, Y.Y. He, X. Jiang, et al., DNA methyltransferase 3B deficiency unveils a new pathological mechanism of pulmonary hypertension, Sci. Adv. 6 (2020) eaba2470.
[29] D. Liu, Y. Yan, J.W. Chen, et al., Hypermethylation of BMPR2 promoter occurs in patients with heritable pulmonary arterial hypertension and inhibits BMPR2 expression, Am. J. Respir. Crit. Care Med. 196 (2017) 925–928.
[30] X.J. Wang, X.Q. Xu, K. Sun, et al., Association of rare PTGIS variants with susceptibility and pulmonary vascular response in patients with idiopathic pulmonary arterial hypertension, JAMA Cardiol. 5 (2020) 677–684.
[31] X.J. Wang, T.Y. Lian, X. Jiang, et al., Germline BMP9 mutation causes idiopathic pulmonary arterial hypertension, Eur. Respir. J. 53 (2019) 1801609.
[32] L. Fu, X. Xu, R. Ren, et al., Modeling xeroderma pigmentosum associated neurological pathologies with patients-derived iPSCs, Protein Cell 7 (2016) 210–221.
[33] G.H. Liu, B.Z. Barkho, S. Ruiz, et al., Recapitulation of premature ageing with iPSCs from Hutchinson-Gilford progeria syndrome, Nature 472 (2018) 221–225.
[34] F. Xie, L. Ye, J.C. Chang, et al., Seamless gene correction of beta thalassemia mutations in patient-specific iPSCs using CRISPR/Cas9 and piggyBac, Genome Res. 24 (2014) 1526–1533.
[35] J.D. Pereira, D.M. DuBreuil, A.C. Devlin, et al., Human sensorimotor organoids derived from healthy and amyotrophic lateral sclerosis stem cells form neuromuscular junctions, Nat. Commun. 12 (2021) 4744.
[36] R.A. Samarasinghe, O.A. Miranda, J.E. Buth, et al., Identification of neural oscillations and epileptiform changes in human brain organoids, Nat. Neurosci. 24 (2021) 1488–1500.
[37] E.A. Imel, Z. Liu, M. Coffman, et al., Oral iron replacement normalizes fibroblast growth factor 23 in iron-deficient patients with autosomal dominant hypophosphatemic rickets, J. Bone Miner. Res. 35 (2020) 231–238.
[38] A. Saeed, J.A. Hoogerland, H. Wessel, et al., Glycogen storage disease type 1a is associated with disturbed vitamin A metabolism and elevated serum retinol levels, Hum. Mol. Genet. 29 (2020) 264–273.
[39] Y. Chen, J. Yu, Y. Niu, et al., Modeling rett syndrome using TALEN-edited MECP2 mutant cynomolgus monkeys, Cell 169 (2017) 945–955 e10.
[40] J. Wang, Z. Wang, H. Zhang, et al., White matter structural and network topological changes underlying the behavioral phenotype of MECP2 mutant monkeys, Cereb. Cortex 31 (2021) 5396–5410.
[41] K. Musunuru, A.C. Chadwick, T. Mizoguchi, et al., *In vivo* CRISPR base editing of PCSK9 durably lowers cholesterol in primates, Nature 593 (2021) 429–434.
[42] H. Yang, L. Shi, B.A. Wang, et al., Generation of genetically modified mice by oocyte injection of androgenetic haploid embryonic stem cells, Cell 149 (2012) 605–617.
[43] L. Wang, Y. Zhang, X. Fu, et al., Joint utilization of genetic analysis and semi-cloning technology reveals a digenic etiology of müllerian anomalies, Cell Res. 30 (2020) 91–94.
[44] Q. Yin, H. Wang, N. Li, et al., Dosage effect of multiple genes accounts for multisystem disorder of myotonic dystrophy type 1, Cell Res. 30 (2020) 133–145.
[45] Y. Lei, R.H. Finnell, New myotonic dystrophy type 1 mouse model, Cell Res. 30 (2020) 99–100.
[46] Y. Yang, Y. Wang, S. Li, et al., Mutations in SCN9A, encoding a sodium channel alpha subunit, in patients with primary erythermalgia, J. Med. Genet. 41 (2004) 171–174.
[47] Z. Lin, Q. Chen, M. Lee, et al., Exome sequencing reveals mutations in TRPV3 as a cause of Olmsted syndrome, Am. J. Hum. Genet. 90 (2012) 558–564.
[48] Z. Song, X. Chen, Q. Zhao, et al., Hair Loss Caused by Gain-of-Function Mutant TRPV3 Is Associated with Premature Differentiation of Follicular Keratinocytes, J. Invest. Dermatol. 141 (2021) 1964–1974.
[49] L. Frésard, C. Smail, N.M. Ferraro, et al., Identification of rare-disease genes using blood transcriptome sequencing and large control cohorts, Nat. Med. 25 (2019) 911–919.
[50] B.B. Cummings, J.L. Marshall, T. Tukiainen, et al., Improving genetic diagnosis in Mendelian disease with transcriptome sequencing, Sci. Transl. Med. 9 (2017) eaal5209.
[51] K.J. Karczewski, M.P. Snyder, Integrative omics for health and disease, Nat. Rev. Genet. 19 (2018) 299–310.
[52] Z. Li, C. Zhu, Y. Ding, et al., ATTEC: a potential new approach to target proteinopathies, Autophagy 16 (2020) 185–187.
[53] Y. Ding, Y. Fei, B. Lu, Emerging new concepts of degrader technologies, Trends Pharmacol. Sci. 41 (2020) 464–474.
[54] Y. Fu, N. Chen, Z. Wang, et al., Degradation of lipid droplets by chimeric autophagy-tethering compounds, Cell Res. 31 (2021) 965–979.
[55] J.R. Mendell, S.A. Al-Zaidy, L.R. Rodino-Klapac, et al., Current clinical applications of in vivo gene therapy with AAVs, Mol. Ther. 29 (2021) 464–488.
[56] L.A. George, S.K. Sullivan, A. Giermasz, et al., Hemophilia B Gene Therapy with a High-Specific-Activity Factor IX Variant, N Engl J Med. 377 (2017) 2215–2227.
[57] Novartis announces landmark EU approval for one-time gene therapy Luxturna® to restore vision in people with rare inherited retinal disease. https://www.novartis.com/news/media-releases/novartis-announces-landmark-eu-approval-one-time-gene-therapy-luxturna-restore-vision-people-rare-inherited-retinal-disease, 2018 (accessed 23 Nov 2018).
[58] Y. Weng, H. Xiao, J. Zhang, et al., RNAi therapeutic and its innovative biotechnological evolution, Biotechnol. Adv. 37 (2019) 801–825.
[59] X. Zhang, L. Wang, M. Liu, et al., CRISPR/Cas9 system: a powerful technology for *in vivo* and ex vivo gene therapy, Sci China Life Sci 60 (2017) 468–475.
[60] D.B. Cox, R.J. Platt, F. Zhang, Therapeutic genome editing: prospects and challenges, Nat. Med. 21 (2015) 121–131.
[61] G.A. Newby, J.S. Yen, K.J. Woodard, et al., Base editing of haematopoietic stem cells rescues sickle cell disease in mice, Nature 595 (2021) 295–302.
[62] Z. Hu, Z. Li, Y. Wu, et al., Targeted B-domain deletion restores F8 function in human endothelial cells and mice, Signal Transduct Target Ther 7 (2022) 189.
[63] CRISPR injected into the blood treats a genetic disease for first time. https://www.science.org/content/article/crispr-injected-blood-treats-genetic-disease-first-time, 2021 (accessed 26 JUN 2021).
[64] D.P. Dever, R.O. Bak, A. Reinisch, et al., CRISPR/Cas9 β-globin gene targeting in human haematopoietic stem cells, Nature 539 (2016) 384–389.
[65] Xiaojie Yan, Qi Pan, Huhu Xin, et al., Genome-editing prodrug: targeted delivery and conditional stabilization of CRISPR-Cas9 for precision therapy of inflammatory disease, Sci. Adv. 7 (2021) eabj0624.
[66] H. Xin, T. Wan, Y. Ping, Off-Targeting of Base Editors: BE3 but not ABE induces substantial off-target single nucleotide variants, Signal Transduct Target Ther 4 (2019) 9.
[67] F. Teng, J. Li, T. Cui, et al., Enhanced mammalian genome editing by new Cas12a orthologs with optimized crRNA scaffolds, Genome Biol. 20 (2019) 15.
[68] F. Teng, T. Cui, G. Feng, et al., Repurposing CRISPR-Cas12b for mammalian genome engineering, Cell Discov 4 (2018) 63.
[69] M.A. Kotterman, T.W. Chalberg, D.V. Schaffer, Viral vectors for gene therapy: translational and clinical outlook, Annu. Rev. Biomed. Eng. 17 (2015) 63–89.
[70] N. Zabaleta, M. Barberia, C. Martin-Higueras, et al., CRISPR/Cas9-mediated glycolate oxidase disruption is an efficacious and safe treatment for primary hyperoxaluria type I, Nat. Commun. 9 (2018) 5454.
[71] X. Wu, L. Chen, X. Wang, Network biomarkers, interaction networks and dynamical network biomarkers in respiratory diseased, Clin. Transl. Med. 3 (2014) 16.
[72] B. Yang, M. Li, W. Tang, et al., Dynamic network biomarker indicates pulmonary metastasis at the tipping point of hepatocellular carcinoma, Nat. Commun. 9 (2018) 678.
[73] C. Li, Y.D. Sun, G.Y. Yu, et al., Integrated omics of metastatic colorectal cancer, Cancer Cell 38 (2020) 734–747.
[74] L. Chen, R. Liu, Z.P. Liu, et al., Detecting early-warning signals for sudden deterioration of complex diseases by dynamical network biomarkers, Sci. Rep. 2 (2013) 342.
[75] X. Liu, X. Chang, S. Leng, et al., Detection for disease tipping points by landscape dynamic network biomarkers, Natl. Sci. Rev. 6 (2019) 775–785.
[76] C. Reily, T.J. Stewart, M.B. Renfrow, et al., Glycosylation in health and disease, Nat. Rev. Nephrol. 15 (2019) 346–366.
[77] Q. Sang, Z. Zhang, J. Shi, et al., A pannexin 1 channelopathy causes human oocyte death, Sci. Transl. Med. 11 (2019) eaav8731.
[78] B. Cheng, Q. Tang, C. Zhang, et al., Glycan Labeling and Analysis in Cells and In Vivo, Annu. Rev. Anal. Chem. (Palo Alto Calif) 14 (2021) 363–387.
[79] D.E. Sun, X. Fan, Y. Shi, et al., Click-ExM enables expansion microscopy for all biomolecules, Nat. Methods 18 (2021) 107–113.

Ruijuan Sun, Ph.D., professor, Executive Deputy Director-General of the Department of Health Sciences, National Natural Science Foundation of China. She has been engaged in medical research and senior scientific administration, contributed to promoting the reformation of funding system, talent cultivation, international cooperation, discipline balance and cross-integration of basic medical research.

Wei Hong, Ph.D., MD, professor. She devoted herself to clinical and basic research of autoimmune diseases for decades, and achieved a series of scientific awards, including the Second Prize of Military Science and Technology Progress Award in 2000 for the mechanistic study on the interaction of fibroblast and extracellular matrix contributing to collagen synthesis in systemic sclerosis. She was employed by the National Natural Science Foundation of China (NSFC) in 2004 and currently serves as Chief in Forth Section, Department of Health Sciences, NSFC. She is now responsible for the management of funds in the field of Medical Immunology, Reproductive system/ Perinatal/ Neonatal Medicine and Medical Genetics in NSFC.