Zhu et al. *Zool. Res.* 2022, 43(2): 234−236
https://doi.org/10.24272/j.issn.2095-8137.2022.068

ZR | Zoological Research

Commentary Open Access

Towards key scientific questions in the diagnosis and treatment of rare diseases: Summary from the 297th Meeting of the Shuangqing Forum

In China, rare diseases are defined as having a birth incidence of less than 1/10 000, or a prevalence of less than 1/10 000 or less than 140 000 patients. Over 7 000 rare diseases affect more than 20 million people in China. Many conditions are misdiagnosed or undiagnosed and most have no treatment, resulting in a huge burden on patients, their families, and the national economy. At the 297th Shuangqing Forum of the National Natural Science Foundation of China, we highlighted the challenges and potential solutions to achieve precision medicine for undiagnosed and rare diseases.

Of the approximately 7 000 rare diseases, more than 80% are considered genetic. Thus far, however, pathogenic variants in genes have only been identified in 20%–40% of these conditions, with most located in protein-coding regions (NIH, 2022; Shire, 2013). Non-protein-coding sequences occupy 98% of the human genome and, depending on their localization and specific interactions with DNA, RNA, and proteins, play key regulatory roles in nearly 93% of diseases (Ding et al., 2021; Hindorff et al., 2009; Liu et al., 2020; Statello et al., 2021). Non-Mendelian inheritance patterns (e.g., oligogenic inheritance), gene-environment interactions, epigenetics, and other chemical modifications of biomolecules (e.g., glycosylation) also warrant further attention to understand the pathogenic mechanisms underlying undiagnosed and rare diseases (Fitz-James & Cavalli, 2022; Kousi & Katsanis, 2015; Oh & Petronis, 2021; Wang et al., 2020). Furthermore, an increasing number of researchers are using systems biology approaches to study the genetic basis of complex diseases, integrating multiple layers representing different scales of biological organization, from the genome to the transcriptome and phenome (Buphamalai et al., 2021; Wong et al., 2021). These integrated multi-level data approaches provide new insights into the characteristics of undiagnosed and rare diseases, thereby facilitating the implementation of next-generation sequencing in clinical diagnosis. In addition, as most existing large-scale studies are limited to European and American enrollment, it is necessary to establish national rare disease cohorts (including longitudinal cohorts) and disease-oriented biobanks in China, such as the GSRD-100K^{WCH} project launched by the West China Hospital of Sichuan University. A national research network is also needed to encourage data sharing and reuse. These innovative research methods, coupled with the increasing availability of population-based data, will increase the discovery rate of causative genes and improve the diagnosis of undiagnosed and rare diseases.

The lack of suitable models for rare diseases has been a barrier to studying disease mechanisms and developing treatment strategies. Human-induced pluripotent stem cells (iPSCs) are well suited for modeling rare diseases as they have the capacity for self-renewal and pluripotency. Furthermore, given the rarity of most monogenic diseases, combining clustered regularly interspaced short palindromic repeats (CRISPR) and iPSCs is a promising approach to determine the pathogenicity of genetic variants (Guo et al., 2021). Gene-edited iPSCs may also be a source of cellular immunotherapy. In combination with high-throughput screening technology, iPSC-derived cells and organoids can be effective tools for screening drug candidates, with implications for the design of therapeutic strategies (Kim et al., 2020; Ryu et al., 2021). At present, animal models do not exist for most rare diseases, and available models are largely limited to rodents due to technical barriers and cost. Furthermore, the gap between rodents and humans is larger than previously understood, especially for diseases related to cognitive function. Therefore, animal models that more closely resemble humans, e.g., non-human primates (Nakamura et al., 2021) and their close relatives (e.g., tree shrews) (Yao, 2017), are needed to evaluate potential causal genes and therapies. Several recent developments, including gene-editing techniques (e.g., CRISPR-Cas) and semi-cloning

Received: 09 February 2022; Accepted: 17 February 2022; Online: 17 February 2022

technology mediated by sperm-like stem cells, have enabled the development of animal disease models predicted to have strong translational relevance (Wang et al., 2020). There is an urgent need to establish a framework for effectively modeling rare diseases at scale using multidisciplinary approaches, which is essential for understanding disease mechanisms and testing prospective therapies.

Rare diseases often exhibit high phenotypic heterogeneity at the early stage, which inevitably affects both diagnosis and treatment. The development of next-generation sequencing technology has made it possible to establish precise clinical diagnosis patterns based on a combination of phenotypes and genotypes, which are especially suitable for rare diseases (The 100 000 Genomes Project Pilot Investigators, 2021; Wright et al., 2018). Novel transformative artificial intelligence/machine learning (AI/ML) strategies and computer automation have been developed to integrate and interpret multi-omics data (Ma et al., 2020). Thus, clinical diagnosis patterns should shift from the classic phenotype-driven model to a genotype-driven sub-molecular model, and precise treatment and disease early warning systems for rare diseases could be established based on both genotypes and phenotypes. The dynamic mapping relationship between genotype and phenotype during disease progression also needs to be further explored (Buphamalai et al., 2021; Wong et al., 2021).

At present, ~263–446 million people worldwide are affected by rare diseases, yet it is estimated that less than 5% of known rare diseases have at least one approved drug for treatment (Nguengang Wakap et al., 2020). Shortages and poor accessibility to orphan drugs remain significant issues in China. Among the 121 rare diseases published in 2018 (first list of rare diseases in China), only 36 have approved pharmacotherapy. There are multiple challenges inherent in developing orphan drugs, including limited understanding of most rare diseases, difficulties in patient enrollment and clinical trial design, and risk of financial unsustainability (Nguengang Wakap et al., 2020). In this context, traditional models of development are inefficient and a new framework is required. In addition to exploring new therapeutic targets, the molecular and pathophysiological mechanisms shared by multiple rare diseases deserve more attention (Tambuyzer et al., 2020). Drug repurposing is an attractive strategy for expanding potential drug candidates for therapy, and greatly reduces the time and cost of drug development (Beijersbergen, 2020). For rare monogenic diseases, gene therapy efficiency can be optimized by improving gene-editing technology and vector delivery systems. More importantly, researchers with expert knowledge in different fields are strongly encouraged to collaborate to develop novel gene and cell therapy products.

Investigations of undiagnosed and rare diseases require multidisciplinary and multisectoral cooperation. Herein, we propose to establish a research alliance, including basic and clinical researchers, institutions, and funding agencies, to promote the development of precision medicine for undiagnosed and rare diseases. Over the next few years, we expect to gain deeper insight into the pathogenic mechanisms and therapeutic strategies of such diseases.

COMPETING INTERESTS

The author declares that he has no competing interests.

AUTHORS' CONTRIBUTIONS

W.H. and C.Y.Z. conceived of the review. C.Y.Z. prepared the draft. All authors contributed to the discussions. All authors read and approved the final version of the manuscript.

Cai-Yun Zhu[2], Wei Hong[1,*], Lei Wang[3], Li-Jun Ding[4], Xue Zhang[5]

[1] *Department of Health Sciences*, *National Natural Science Foundation of China*, *Beijing* 100085, *China*
[2] *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*
[3] *Institute of Pediatrics*, *Children's Hospital of Fudan University and Shanghai Key Laboratory of Medical Epigenetics*, *International Co-laboratory of Medical Epigenetics and Metabolism*, *Ministry of Science and Technology and Institutes of Biomedical Sciences*, *State Key Laboratory of Genetic Engineering*, *Fudan University*, *Shanghai* 200032, *China*
[4] *Center for Reproductive Medicine and Obstetrics and Gynecology*, *Affiliated Drum Tower Hospital of Nanjing University Medical School*, *Nanjing*, *Jiangsu* 210008, *China*
[5] *McKusick-Zhang Center for Genetic Medicine*, *Institute of Basic Medical Sciences*, *Chinese Academy of Medical Sciences/School of Basic Medicine*, *Peking Union Medical College*, *Beijing* 100005, *China*
*Corresponding author, E-mail: hongwei@nsfc.gov.cn

REFERENCES

Beijersbergen RL. 2020. Old drugs with new tricks. *Nature Cancer*, **1**(2): 153−155.

Buphamalai P, Kokotovic T, Nagy V, Menche J. 2021. Network analysis reveals rare disease signatures across multiple levels of biological organization. *Nature Communications*, **12**(1): 6306.

Ding B, Liu Y, Liu ZH, Zheng LN, Xu P, Chen Z, et al. 2021. Noncoding loci without epigenomic signals can be essential for maintaining global chromatin organization and cell viability. *Science Advances*, **7**(45): eabi6020.

Fitz-James MH, Cavalli G. 2022. Molecular mechanisms of transgenerational epigenetic inheritance. *Nature Reviews Genetics*,doi: 10.1038/s41576-021-00438-5.

Guo HC, Liu LC, Nishiga M, Cong L, Wu JC. 2021. Deciphering pathogenicity of variants of uncertain significance with CRISPR-edited iPSCs. *Trends in Genetics*, **37**(12): 1109−1123.

Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, et al. 2009. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proceedings of the National Academy of Sciences of the United States of America*, **106**(23): 9362−9367.

Kim J, Koo BK, Knoblich JA. 2020. Human organoids: model systems for human biology and medicine. *Nature Reviews Molecular Cell Biology*, **21**(10): 571−584.

Kousi M, Katsanis N. 2015. Genetic modifiers and oligogenic inheritance. *Cold Spring Harbor Perspectives in Medicine*, **5**(6): a017145.

Liu Y, Li CL, Shen SH, Chen XL, Szlachta K, Edmonson MN, et al. 2020. Discovery of regulatory noncoding variants in individual cancer genomes by using cis-X. *Nature Genetics*, **52**(8): 811−818.

Ma AJ, McDermaid A, Xu J, Chang YZ, Ma Q. 2020. Integrative methods and practical challenges for single-cell multi-omics. *Trends in Biotechnology*, **38**(9): 1007−1022.

Nakamura T, Fujiwara K, Saitou M, Tsukiyama T. 2021. Non-human primates as a model for human development. *Stem Cell Reports*, **16**(5): 1093−1103.

Nguengang Wakap S, Lambert DM, Olry A, Rodwell C, Gueydan C, Lanneau V, et al. 2020. Estimating cumulative point prevalence of rare diseases: analysis of the Orphanet database. *European Journal of Human Genetics*, **28**(2): 165−173.

NIH. 2022. NIH genetic and rare disease information center (GARD) FAQs about rare disease. https://rarediseases.info.nih.gov/diseases/pages/31/faqs-about-rare-diseases/.

Oh ES, Petronis A. 2021. Origins of human disease: the chrono-epigenetic perspective. *Nature Reviews Genetics*, **22**(8): 533−546.

Ryu S, Chu PH, Malley C, Braisted J, Ormanoglu P, Huang RL, et al. 2021. Human pluripotent stem cells for high-throughput drug screening and characterization of small molecules. *In*: Walker JM (Eds). Methods in Molecular Biology. New York: Springer.

Shire. 2013. Rare disease impact report: insights from patients and the medical community.https://globalgenes.org/wp-content/uploads/2013/04/ShireReport-1.pdf.

Statello L, Guo CJ, Chen LL, Huarte M. 2021. Gene regulation by long non-coding RNAs and its biological functions. *Nature Reviews Molecular Cell Biology*, **22**(2): 96−118.

Tambuyzer E, Vandendriessche B, Austin CP, Brooks PJ, Larsson K, Miller Needleman KI, et al. 2020. Therapies for rare diseases: therapeutic modalities, progress and challenges ahead. *Nature Reviews Drug Discovery*, **19**(2): 93−111.

The 100, 000 Genomes Project Pilot Investigators. 2021. 100, 000 genomes pilot on rare-disease diagnosis in health care—preliminary report. *The New England Journal of Medicine*, **385**(20): 1868−1880.

Wang LB, Zhang Y, Fu XY, Dong SS, Tang SY, Zhang N, et al. 2020. Joint utilization of genetic analysis and semi-cloning technology reveals a digenic etiology of Müllerian anomalies. *Cell Research*, **30**(1): 91−94.

Wong AK, Sealfon RSG, Theesfeld CL, Troyanskaya OG. 2021. Decoding disease: from genomes to networks to phenotypes. *Nature Reviews Genetics*, **22**(12): 774−790.

Wright CF, FitzPatrick DR, Firth HV. 2018. Paediatric genomics: diagnosing rare disease in children. *Nature Reviews Genetics*, **19**(5): 253−268.

Yao YG. 2017. Creating animal models, why not use the Chinese tree shrew (*Tupaia belangeri chinensis*)?. *Zoological Research*, **38**(3): 118−126.