A Diagnosis for All Rare Genetic Diseases: The Horizon and the Next Frontiers

Kym M. Boycott,[1,*] Taila Hartley,[1] Leslie G. Biesecker,[2] Richard A. Gibbs,[3] A. Micheil Innes,[4] Olaf Riess,[5] John Belmont,[6,2] Sally L. Dunwoodie,[7] Nebojsa Jojic,[8] Timo Lassmann,[9] Deborah Mackay,[10] I. Karen Temple,[10] Axel Visel,[11,12,13] and Gareth Baynam[14,15,16]

[1]Children's Hospital of Eastern Ontario Research Institute, University of Ottawa, Ottawa, ON, Canada
[2]Medical Genomics and Metabolic Genetics Branch, National Human Genome Research Institute, Bethesda, MD, USA
[3]Human Genome Sequencing Center, Department of Human and Molecular Genetics, Baylor College of Medicine, Houston, TX, USA
[4]Department of Medical Genetics and Alberta Children's Hospital Research Institute, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada
[5]Institute of Medical Genetics and Applied Genomics, University of Tübingen, Tübingen, Germany
[6]Illumina, Madison, WI, USA
[7]Victor Chang Cardiac Research Institute, Faculty of Medicine, University of New South Wales, Sydney, NSW, Australia
[8]Microsoft Research, Seattle, Washington, USA
[9]Telethon Kids Institute, University of Western Australia, Nedlands, WA, Australia
[10]Department of Human Genetics and Genomic Medicine, Faculty of Medicine, University of Southampton, Southampton, Hampshire, UK
[11]Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, CA, USA
[12]DOE Joint Genome Institute, CA, USA
[13]The University of California at Merced, CA, USA
[14]Faculty of Health and Medical Sciences, University of Western Australia Medical School, Perth, WA, Australia
[15]Western Australian Register of Developmental Anomalies, Genetic Services of Western Australia, Perth, WA, Australia
[16]Office of Population Health Genomics, Western Australian Department of Health, Perth, WA, Australia
*Correspondence: kboycott@cheo.on.ca
https://doi.org/10.1016/j.cell.2019.02.040

The introduction of exome sequencing in the clinic has sparked tremendous optimism for the future of rare disease diagnosis, and there is exciting opportunity to further leverage these advances. To provide diagnostic clarity to all of these patients, however, there is a critical need for the field to develop and implement strategies to understand the mechanisms underlying all rare diseases and translate these to clinical care.

Introduction

Hundreds of millions of lives are affected by an estimated 10,000 unique genetically determined diseases. Individually, each disease affects a relatively small number of people, leading to their common label as rare genetic diseases (RDs); however, collectively, they represent an important public health opportunity. The vast majority of these patients experience long and grueling diagnostic odysseys and lack treatment. In 2011, recognition of both the longstanding inequity in care and the great opportunity for tractability due to technical developments led to the founding of the International Rare Diseases Research Consortium (IRDiRC), which aims to advance global cooperation among numerous stakeholders (Dawkins et al., 2018). The vision of IRDiRC is to enable all people living with a RD to receive an accurate diagnosis, care, and available therapy within 1 year of coming to medical attention (Austin et al., 2018).

Achieving an accurate and timely molecular diagnosis will largely depend on progress in the discovery of the genes and genetic mechanisms associated with RDs. While the exact number of RDs is debated (Hartley et al., 2018), it is estimated that thousands of RD genes and disease mechanisms remain undiscovered. Over the past 8 years, exome sequencing (ES) in both research and clinical settings has been a powerful tool for discovering new disease genes for RDs that were intractable to previous approaches. Most advances have been for highly recognizable clinical presentations associated with early age of onset and significant morbidity and mortality and caused by highly penetrant (typically protein-coding) variants (Boycott et al., 2017). The diagnostic utility of ES has translated beautifully into the clinic, with a diagnostic yield in the range of 25%–30% among large and heterogeneous RD cohorts (Clark et al., 2018). Here, we

discuss the continued importance of ES in both the clinic and the research environment, the next wave of technologies on the horizon, and the next frontiers for RD discovery, moving toward the ultimate goal of diagnostic clarity for each and every family affected by a RD.

Achieving a Diagnosis for All
The "Here and Now": The Continued Role of Exome Sequencing

The application of ES for RD patients represents a remarkable achievement in diagnostics, with a diagnostic yield far higher than other genetic tests (Clark et al., 2018). Nonetheless, in >70% of patients in whom there was a high degree of pre-test suspicion for a monogenic RD, ES provides no molecular diagnosis. For the benefit of RD patients, it is imperative that we drive this diagnostic yield to as close to 100% as possible. While the theoretical yield of ES is unknown, in patient populations with specific

presentations and a high degree of certainty that there is a genetic cause to the RD, the yield of the coding genome is likely well over 50% (Beaulieu et al., 2014; Shamseldin et al., 2017). Indeed, there remains substantial diagnostic potential in existing ES data. For starters, evidence is emerging that reanalysis of negative clinical ES data just 1 to 3 years later increases diagnostic yield by 10% (Wenger et al., 2017). This is because at initial analysis, there was insufficient evidence for candidate variant or gene causality, but this evidence emerges upon reanalysis in light of the annual curation of >10,000 disease variants (ClinVar [https://www.ncbi.nlm.nih.gov/clinvar/] and HGMD [http://www.hgmd.cf.ac.uk/ac/index.php]) and 250 novel disease-gene associations (OMIM [https://www.omim.org/] and Orphanet [https://www.orpha.net/consor/cgi-bin/index.php]). Even higher diagnostic yields can be achieved through reanalysis in collaboration with the referring physician, with estimates as high as 12% (Salmon et al., 2018). Collaboration with research laboratories can provide additional increases (Eldomery et al., 2017) boosted by the application of novel computational tools, sequencing of additional family members, and gene-discovery efforts. These strategies have been bolstered by platforms that share genotype and phenotype information to identify patients with overlapping phenotypes and candidate genes, an approach called matchmaking (reviewed in Philippakis et al., 2015; www.matchmakerexchange.org). While technical limitations of ES are well recognized, in the last decade, capture kits have continued to enhance their coverage of the coding genome, with additional features to provide coverage of previously reported variants in promoters and deep intronic regions of known disease genes. In addition, computational tools continue to improve and facilitate the identification of variation. Given cost and other practical considerations, ES will continue to play a major role in RD variant diagnosis and discovery.

Globally, thousands of clinical exomes are performed weekly, but unfortunately, the majority of these data are inaccessible for discovery and matchmaking. To realize the theoretical maximal diagnostic yield of ES will require a globally coordi-

nated paradigm shift; every patient must have the opportunity to be a research patient. More international and less restrictive data sharing is critical to drive disease gene discovery, facilitate variant interpretation, enhance control datasets, and develop new computational tools. This will enable identification of RDs that are understood at a genetic level, while RDs that require further research can be studied as part of an "exome-negative" clinical infrastructure. Importantly, this paradigm shift will only be realized if, in parallel, we develop the appropriate computational architecture, ensure protection of participant privacy, and continue to promote the cultural shifts that will enable data sharing on a global scale. Importantly, aggregation of such data will contribute to the development of large datasets that can be used for as-yet undefined purposes as we explore new mechanisms for RD.

The "Horizon" in RD Diagnosis: The Next Wave of Technologies to Reveal RD Mechanisms

Regardless of the ultimate capability of ES to provide diagnoses for RD patients, some disease mechanisms are difficult or impossible to detect using this approach (Table 1). For example, mosaicism of a pathogenic variant would not be routinely identified by current analytical approaches. Challenges for detection of mosaicism include the distribution of the causative genomic variation, which can be non-random and can exclude the most often-sampled tissue (blood) for genetic testing, the changing level of mosaicism over time, the difficulty in distinguishing pathogenic from benign or unrelated mosaic variation (signal to noise), and the high sequencing cost for the depth and breadth of coverage needed to detect low-level mosaic variants. New data-analysis tools are emerging to screen for mosaicism in unsolved exome datasets, and approaches that facilitate very deep sequencing of targeted regions in a cost-effective manner will improve detection of mosaicism in the near term.

Some pathogenic genomic variants are missed entirely by ES. Genome sequencing (GS) (short-read sequencing) outperforms ES for indels (small insertion-deletions), copy-number variations (CNVs), and chromosomal rearrange-

ments, while long-read GS promises further improvements in detection of rearrangements and the ability to identify RDs secondary to pathogenic repeat expansions. GS also provides the opportunity to identify regulatory variants that lie outside the exome, such as in promoters, enhancers, deep intronic regions, or distant-acting regulatory sequences located in intergenic regions, though interpretation of such variants and proof of causality are challenging. Such advantages of GS are the basis for promoting this approach over ES, and while robust head-to-head comparisons of the two approaches are still lacking, we hypothesize that GS will increase the diagnostic yield of a genome-wide clinical test by at least 10% in the near term. As clinical GS data accumulate and understanding of intronic and intergenic variation improves, this yield will significantly increase over the years.

Several emerging technologies offer value as adjunct diagnostic tools by providing an approach to assess the functional significance of variants. For example, transcriptome sequencing can evaluate the functional consequences of variants that may affect splicing or gene regulation (e.g., decreased, increased, or monoallelic gene expression). This approach has been suggested to increase the diagnostic yield by 10%–35% in known genes for certain clinical indications. Although promising, its broad applicability for RDs is unknown given challenges around the availability of relevant tissues, including those at critical stages of development. Similarly, methylation arrays are providing functional insight into imprinting disorders, which are caused by alterations of the expressed copy number of imprinted genes, through epigenetic error, uniparental disomy, or CNVs/single-nucleotide variants (SNVs) of the regulatory DNA or the expressed allele (Soellner et al., 2017). More than 100 human germline-imprinted genes distributed across the genome have been identified, and it is likely that more remain to be found. In addition, arrays can detect specific DNA methylation epi-signatures for RDs associated with chromatin dysregulation; these syndrome-specific biomarkers complement standard clinical diagnostics (Aref-Eshghi et al., 2018). The true prevalence and phenotypic

Table 1. Mechanisms of RD that are Currently Intractable to Exome Sequencing and Analysis

| Mechanism			Description	Approaches
General Mechanism	Perspective for Solutions[a]	Mechanism Subcategory		
Mosaicism	horizon	tissue-specific mosaicism	mosaic manifestations of Mendelian disorders; disorders that manifest only as mosaicism	deep sequencing of multiple tissues
Genomic alterations	horizon	small insertions/deletions	small structural changes missed by ES and microarray (<50 bp)	GS
		large insertions/deletions	larger structural changes missed by ES and microarray (>50bp)	GS
		chromosomal rearrangements	inversions/translocations; multiple deletions/duplications	GS
		repeat expansions	triplet and other expansions	long-read GS
		transposable elements (retrotransposons)	genomic sequences that copy and paste into locations throughout the genome (such as mobile element insertions)	novel approaches to data analysis
Gene regulation	horizon	splicing mutations	synonymous or splice site or intronic mutations	GS, RNA sequencing (RNA-seq)
		imprinting	altered parent-of-origin specific expression pattern	methylation arrays
	next frontier	regulatory DNA mutations	promoter, enhancer, and other regulatory mutations	GS, RNA-seq, High-C, prediction tools
		noncoding RNA mutations	intronic, intergenic, and antisense RNAs (e.g., microRNAs, small nucleolar RNAs [snoRNAs])	novel approaches to data analysis
		mutations that alter post-transcriptional or post-translational modifications	altered RNA or protein modifications that impact stability or catalytic function	novel approaches to data analysis
Complex inheritance	horizon	unusual or less common inheritance patterns	sex-limited expression, paradoxical inheritance, necessary but not sufficient CNVs, uniparental disomy	novel approaches to data analysis
	next frontier	genetic modifiers	allele from one gene reduces or exacerbates the penetrance or expressivity of phenotype associated with another gene	novel approaches to data analysis; validation in model organisms
		gene-environment interaction	rare susceptibility allele combined with environmental trigger	environmental exposure data capture, validation in model organisms, metabolomics
		maternal effects	mutation in the mother results in altered fetal environment	environmental-exposure data capture; validation in model organisms
		digenic, oligogenic, or polygenic	interaction of two or more genes	novel approaches to data analysis; validation in model organisms

[a]Horizon: near-term (within 5 years); frontier: longer-term (5 years and beyond).

spectra of imprinting disorders will only be determined by the coordinated implementation of genomic and epigenomic technologies and recognition that the right family member to analyze might not be the affected individual. Similarly, atypical inheritance patterns should be considered when analyzing genomic data of unsolved RD patients; this will require even more sophisticated approaches to data analysis that will identify such mechanisms in a diagnostic setting (Table 1). Finally, for all of these new technologies, diagnostic standards will need to be established before clinical implementation to facilitate diagnostic clarity for as many patients as possible.

The "Next Frontiers" in RD Discovery: Building out from Mendelian Inheritance

Despite the excitement around GS, few RD discoveries have been made outside of the protein-coding regions of the genome. Comprehensive analysis of the noncoding genome on a broader scale represents a significant frontier (Table 1).
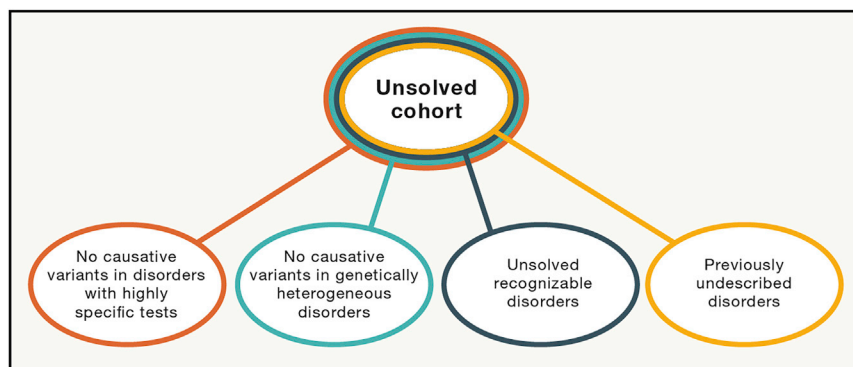
Figure 1. Clinical Groupings of the Unsolved RD Cohort
The unsolved cohort of patients can be considered in four groups; each will require a multifaceted approach and will give us different insights into the incredible landscape of mechanisms underlying RD.

The opportunity lies in the interpretation of noncoding variation, which is exponentially more difficult given the unresolved complexity of how noncoding DNA regulates gene expression, lack of adequate control datasets, and computational tools to predict variant impact and the fact that each of these noncoding variants is likely affecting only a single patient or family, resulting in a high benchmark to establish pathogenicity. While confirming predicted splicing abnormalities is relatively straightforward, as highlighted in the previous section, estimating the impact of mechanisms such as long-range DNA regulation, aberrant DNA modifications (such as methylation), pathogenic alterations to non-coding RNA, and post-transcriptional and post-translational dysregulation on RD will require significantly greater understanding of the genome and major advances in functional analytical approaches. Initial successes will likely center on the use of large families with linkage data to narrow the search space and a focus on noncoding *de novo* alterations in parent-affected child studies.

Alongside monogenic RDs, we will face the challenge of RDs of complex etiology, with a primary genetic driver but clinical presentations that are contextualized by additional factors and these RDs represent another significant frontier of study. The relative impact of genetic and environmental components on RDs will depend on the underlying mechanism of interaction (signal transduction pathway, unfolded protein response, epigenetic modifications, etc.) and in the case of

embryogenesis, when during development the impact is elicited and which developing organs/tissues are most vulnerable to perturbation at a given time. Environmental exposures may be pre- or post-natal, and the challenge will be to capture such exposure information based on history as well as dynamic biological data. Recently, insight into the thousands of metabolic reactions occurring within the human body (e.g., the metabolome) has shown promise as a readout of genes and the environment at a particular point in time. Such studies will require integration of epidemiological and multi-omic data in exposed and non-exposed populations. Experimental support for such complex mechanisms will require the use of functional assays and model organisms to validate findings (Shi et al., 2017).

By comparison, the investigation of digenic, oligogenic, and polygenic inheritance models may seem relatively straightforward, but one should not be deceived, and this represents yet another frontier. To perform such analyses and collect the evidence required for the statistical certainty needed to support an RD mechanism, a massive amount of harmonized phenotypic, genotypic, and family history data will be required. The establishment of such datasets reinforces the need to offer research access and broad data sharing to all RD patients and their families.

The Unsolved Rd Cohort: The Way Forward
Our ability to diagnose all RDs is limited by our incomplete understanding of the full

mutational spectrum associated with all RDs and the sheer number of unique RDs that have yet to be defined. The way forward is readily recognized as multifaceted and will likely focus on specific subsets of patients from the unsolved RD cohort (Figure 1); each subset has significant utility for exploring RD mechanisms and optimizing approaches for clinical translation of novel diagnostic tests. Patients in the unsolved RD cohort can be considered in four groups, and while the approaches used to uncover the genetic mechanism for the respective RDs may be similar between groups, the knowledge gained for each will be unique.

Patients with No Causative Variant after an Appropriate, Highly Sensitive Test
Patients in this group have had the appropriate genetic test that is highly sensitive for that particular RD but remain without a molecular diagnosis (e.g., single-gene disorders such as cystic fibrosis and neurofibromatosis type 1). In all likelihood, the causative variant(s) in these patients is/are not detected by the current testing methodologies, and therefore, this subset of patients represents a remarkable opportunity to explore novel diagnostic approaches, including new technologies and computational tools, to more comprehensively assess the spectrum of possible genetic causes of a given disease. The insights delivered will be directly relevant to these patients while also optimizing patient sampling, computational tools, and diagnostic algorithms based on emerging technologies. More broadly, such knowledge will contribute significantly to the mechanistic spectrum of other unsolved RDs. This type of exploratory focus represents a shift in the types of studies that our community traditionally values, and both funders and publishers will need to recognize the intermediate importance and long-term impacts of the resulting insights.

Patients with No Identified Causative Variant in the Context of Genetic Heterogeneity
Patients in this group have a clinically recognizable presentation associated with genetic heterogeneity (e.g., hereditary spastic paraplegia, myopathy, retinitis pigmentosa) but negative results for the appropriate testing and analysis, including most of the relevant disease

genes. These patients have either a pathogenic variant in one of the known disease genes that was not detected using the current testing approach or a yet-to-be-discovered disease-associated gene. To diagnose these patients, we will need large datasets of patients that include detailed phenotypic and genomic information for data comparison and novel technologies and computational tools to identify cryptic variants. Besides GS, transcriptomic, metabolomic, epigenomic, and proteomic data may be necessary to identify the underlying genetic causes, particularly for simplex patients in which family-based analyses are not possible. This patient group represents an opportunity similar to the subset described above but is also enriched for novel disease-gene discovery and is likely one of the largest populations in the unsolved cohort.

Patients with an Unsolved but Recognizable Syndrome

Patients in this group have a clinical diagnosis based on similarity to a previously described syndrome for which the underlying etiology is unknown (e.g., PHACE and Hallermann-Streiff syndromes, VACTERL association). With the efforts of the clinical and scientific communities and the increased use of ES, we now understand the genetic basis of most of the frequent and recognizable human malformation syndromes. However, some well-established syndromes (defined as reported in >10 unrelated patients and curated by OMIM and Orphanet) remain without an understood molecular etiology despite intensive investigation. Examples of such unsolved syndromes have been recently reviewed (Boycott et. al., 2018) in a special issue on Unsolved Recognizable Patterns of Human Malformation: Challenges and Opportunities in the American Journal of Medical Genetics. Possible explanations for their current intractability include genetic and phenotypic heterogeneity, mosaicism, epigenetics, gene-environment interactions, and other non-Mendelian contributions. The way forward for this group of disorders will require the use of emerging and new technologies, global cooperation, and data sharing.

Patients with Syndromes without a Name

The fourth group of patients present with a constellation of clinical symptoms and signs that are not recognizable as a previously described syndrome or condition, which may have non-specific clinical features, and fit into none of the above groups. Part of the challenge in the diagnosis of these patients is that the full extent of the clinical presentation may not yet have become manifest (as occurs in the evaluation of ill newborns). These patients are most suitable for genome-wide sequencing approaches and, for the foreseeable future, should undergo ES or GS as a first-line test followed by detailed genotypic and phenotypic data sharing for matchmaking purposes. Their diagnoses will likely include early presentations of recognized RDs, expanded phenotypes of previously recognized RDs, and novel RDs associated with new genes that will only be identified once RD datasets contain sufficient genotypic and phenotypic data to provide statistical confidence that an accurate diagnosis has been made and/or following validation in model organisms.

Conclusions

We face a grand opportunity in precision public health: to understand the cause of each and every RD and provide a diagnosis for each individual RD patient. Clinical ES is transforming molecular diagnosis and will continue to have a remarkable impact on this area of medicine. For the patients that remain unsolved after genetic testing, the future remains optimistic. A large number of emerging technologies are on the horizon and will play an important role in RD diagnostics in the near term. Computational approaches that focus on large-scale data integration across patients and within the single patient ("systems diagnostics"), and from healthy individuals, will enable the next frontiers of RD discovery. As we work toward our goal of diagnostic clarity for all, we will gain important insights into the RD genome and the attendant knowledge about human biology that this will bring. Importantly, there are some cross-cutting requisites for the clinical and research community to enable this important work and reach not just the current horizons of RD diagnostics, but the next frontiers as well. To start, we need to provide all RD patients the ability to access clinical genome-wide testing and participate in research.

At the health-systems level, we must implement the timely, prioritized, and sustainable clinical integration of proven innovative diagnostic approaches; this will scale the input side of the equation, serving patient needs and fueling translational research discovery. In facilitating research participation, we must include those that we do not typically consider, collecting data for those with molecular diagnoses, and the clinically diagnosed but causative-variant-negative patients, and support the necessary infrastructure. Going forward, we need to address the fundamental lack of RD researchers that study complex mechanisms by enabling an emerging new generation of scientists in this area with adequate funding and by contributing to comprehensive RD datasets that will provide the foundation for their work. Most critically, we must recognize that the future of RD diagnostics will depend on the international RD community working as one team toward an ambitious and important joint goal. We need to overcome a mindset limited to individual patients; individual researchers; individual genetic mechanisms; and even individual consortia, countries, or continents. The vision of IRDiRC, for each RD patient to receive a diagnosis within 1 year, is achievable only if we collectively take up this grand opportunity on a global scale.

DECLARATION OF INTEREST

J.B. is a full time employee of Illumina, Inc. L.G.B. is a member of the Illumina Medical Ethics committee, receives royalties from Genentech, and receives honoraria from Cold Spring Harbor Press. The remaining authors have no conflict of interest to declare.

REFERENCES

Aref-Eshghi, E., Rodenhiser, D.I., Schenkel, L.C., Lin, H., Skinner, C., Ainsworth, P., Paré, G., Hood, R.L., Bulman, D.E., Kernohan, K.D., et al.; Care4Rare Canada Consortium (2018). Genomic DNA Methylation Signatures Enable Concurrent Diagnosis and Clinical Genetic Variant Classification in Neurodevelopmental Syndromes. Am. J. Hum. Genet. 102, 156–174.

Austin, C.P., Cutillo, C.M., Lau, L.P.L., Jonker, A.H., Rath, A., Julkowska, D., Thomson, D., Terry, S.F., de Montleau, B., Ardigò, D., et al.; International Rare Diseases Research Consortium (IRDiRC) (2018). Future of Rare Diseases Research 2017-2027: An IRDiRC Perspective. Clin. Transl. Sci. 11, 21–27.

Beaulieu, C.L., Majewski, J., Schwartzentruber, J., Samuels, M.E., Fernandez, B.A., Bernier, F.P., Brudno, M., Knoppers, B., Marcadier, J., Dyment, D., et al.; FORGE Canada Consortium (2014). FORGE Canada Consortium: outcomes of a 2-year national rare-disease gene-discovery project. Am. J. Hum. Genet. 94, 809–817.

Boycott, K.M., Rath, A., Chong, J.X., Hartley, T., Alkuraya, F.S., Baynam, G., Brookes, A.J., Brudno, M., Carracedo, A., den Dunnen, J.T., et al. (2017). International Cooperation to Enable the Diagnosis of All Rare Genetic Diseases. Am. J. Hum. Genet. 100, 695–705.

Boycott, K.M., Dyment, D.A., and Innes, A.M. (2018). Unsolved recognizable patterns of human malformation: Challenges and opportunities. Am. J. Med. Genet. C. Semin. Med. Genet. 178, 382–386.

Clark, M.M., Stark, Z., Farnaes, L., Tan, T.Y., White, S.M., Dimmock, D., and Kingsmore, S.F. (2018). Meta-analysis of the diagnostic and clinical utility of genome and exome sequencing and chromosomal microarray in children with suspected genetic diseases. NPJ Genom. Med. 3, 16.

Dawkins, H.J.S., Draghia-Akli, R., Lasko, P., Lau, L.P.L., Jonker, A.H., Cutillo, C.M., Rath, A., Boycott, K.M., Baynam, G., Lochmüller, H., et al.; International Rare Diseases Research Consortium (IRDiRC) (2018). Progress in Rare Diseases Research 2010-2016: An IRDiRC Perspective. Clin. Transl. Sci. 11, 11–20.

Eldomery, M.K., Coban-Akdemir, Z., Harel, T., Rosenfeld, J.A., Gambin, T., Stray-Pedersen, A., Küry, S., Mercier, S., Lessel, D., Denecke, J., et al. (2017). Lessons learned from additional research analyses of unsolved clinical exome cases. Genome Med. 9, 26.

Hartley, T., Balcı, T.B., Rojas, S.K., Eaton, A., Canada, C.R., Dyment, D.A., and Boycott, K.M. (2018). The unsolved rare genetic disease atlas? An analysis of the unexplained phenotypic descriptions in OMIM®. Am. J. Med. Genet. C. Semin. Med. Genet. 178, 458–463.

Philippakis, A.A., Azzariti, D.R., Beltran, S., Brookes, A.J., Brownstein, C.A., Brudno, M., Brunner, H.G., Buske, O.J., Carey, K., Doll, C., et al. (2015). The Matchmaker Exchange: a platform for rare disease gene discovery. Hum. Mutat. 36, 915–921.

Salmon, L.B., Orenstein, N., Markus-Bustani, K., Ruhrman-Shahar, N., Kilim, Y., Magal, N., Hubshman, M.W., and Bazak, L. (2018). Improved diagnostics by exome sequencing following raw data reevaluation by clinical geneticists involved in the medical care of the individuals tested. Genet. Med..

Shamseldin, H.E., Maddirevula, S., Faqeih, E., Ibrahim, N., Hashem, M., Shaheen, R., and Alkuraya, F.S. (2017). Increasing the sensitivity of clinical exome sequencing through improved filtration strategy. Genet. Med. 19, 593–598.

Shi, H., Enriquez, A., Rapadas, M., Martin, E.M.M.A., Wang, R., Moreau, J., Lim, C.K., Szot, J.O., Ip, E., Hughes, J.N., et al. (2017). NAD Deficiency, Congenital Malformations, and Niacin Supplementation. N. Engl. J. Med. 377, 544–552.

Soellner, L., Begemann, M., Mackay, D.J., Grønskov, K., Tümer, Z., Maher, E.R., Temple, I.K., Monk, D., Riccio, A., Linglart, A., et al. (2017). Recent Advances in Imprinting Disorders. Clin. Genet. 91, 3–13.

Wenger, A.M., Guturu, H., Bernstein, J.A., and Bejerano, G. (2017). Systematic reanalysis of clinical exome data yields additional diagnoses: implications for providers. Genet. Med. 19, 209–214.