

Analysis of Genes Associated with Rare Genetic Diseases “An Applied Study Using Genomic Techniques in Local Communities”

Shifaa Jawad jaaz

Ministry of Education, Dhi Qar Education Directorate

Received: 2024, 15, Dec

Accepted: 2024, 21, Dec

Published: 2025, 08, Jan

Copyright © 2025 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).



Open Access

<http://creativecommons.org/licenses/by/4.0/>

Annotation: This study highlights the transformative role of genomics in diagnosing and understanding rare genetic diseases (RGDs), which pose a medical and research challenge due to their complexity and diversity. Using whole exome sequencing (WES), genes associated with these diseases were analyzed in local communities, resulting in a confirmed or probable diagnosis in 29.2% of the cases studied.

The study showed that trio analysis achieved the highest diagnosis rates (40%), while neurodevelopmental diseases and systemic disorders benefited more from these techniques than isolated patterns. However, the results revealed ongoing challenges in interpreting genetic variants, calling for updating databases and adopting technologies such as whole genome sequencing (WGS) to complete the analysis.

The study highlights the importance of integrating the genomic approach into local healthcare systems, pointing to the potential benefits of personalized medicine and community healthcare, especially in resource-limited areas. Local discoveries

also highlighted novel genetic variants contributing to a deeper understanding of disease mechanisms.

The results indicate the need for enhanced international collaboration and continued longitudinal research to ensure the effective and sustainable application of genomic medicine technologies in different communities.

Keywords: Rare diseases, genetic analysis, genomic techniques,

Introduction

Despite their rarity, rare genetic diseases collectively impact a significant portion of the global population. These illnesses are frequently brought on by particular genetic mutations that interfere with regular biological functions. Advances in genomic technologies, particularly whole genome sequencing (WGS) and whole exome sequencing (WES), have revolutionized the ability to identify such mutations. However, the application of these technologies in local communities remains limited, especially in areas with limited access to genetic healthcare. This project focuses on applying genomic technology to evaluate genes linked with uncommon genetic illnesses in local communities, emphasizing community participation and translational effect.

Despite affecting 1 in 16 persons globally and accounting for over 9,000 instances, uncommon diseases pose a substantial cost to society and cause many people to suffer from life-limiting conditions. Although the precise number is unknown, it is estimated that 80% of uncommon diseases have a genetic basis and that between 57 and 67% of them shorten life expectancy. Rare genetic disorders account for a large portion of healthcare expenditures, with around 39% of rare diseases having a known genetic basis. Compared to typical chronic diseases, rare genetic disorders are substantially more expensive. Fifty percent of patients who are suspected of having a rare genetic condition never acquire a molecular diagnosis, even though they frequently undergo lengthy diagnostic processes that take more than five years and involve several visits and specialized testing. [1]

Even though there aren't many authorized therapies for rare genetic disorders, a precise genetic diagnosis is essential for care and family well-being, including emotional support, planning, and counseling. In the upcoming ten years, new methods provide undiagnosed patients hope of learning about new illness causes and receiving a diagnosis.

Both whole exome and whole genome sequencing are effective ways to find genetic diversity. Because whole exome sequencing focuses on the protein-coding region of the genome and is inexpensive, it is the most often used approach for identifying the genes that cause uncommon disorders. Whole exome sequencing involves enriching and sequencing the exome, which makes up around 1% of the genome, and then analyzing the results. Finding the genes causing disorders like Freeman-Sheldon syndrome and Miller syndrome was one of the first successes of whole exome sequencing in rare illness research. Since then, whole exome sequencing has been used to identify over 180 additional genes, but the best methods for doing so are still being worked out.

Rare diseases encompass a broad spectrum of illnesses with various etiology processes and are

characterized by their low prevalence. Less than 200,000 individuals in the US are afflicted by an illness, which includes 6,000 to 10,000 uncommon diseases that impact millions of people globally. Although non-genetic uncommon diseases like some malignancies and autoimmune disorders also exist, genetic rare diseases make up around 80% of these illnesses. Researching uncommon genetic illnesses offers the chance to learn novel biology that may significantly affect prevalent illnesses. Significant scientific discoveries that have earned Nobel Prizes have been achieved via the study of uncommon genetic variations.[1]

These findings advance our knowledge of how rare metals affect common neurological conditions like Menkes disease. This highlights the significance of researching rare genetic diseases to comprehend disease mechanisms more generally, which is what we will cover in this study as we use genomic technologies to analyze the genes linked to rare genetic diseases. [1-4]

	<i>Number of genes discovered</i>	<i>notes</i>
2018	30	The beginning of the widespread use of WES in genetic diagnosis
2019	45	Using AI to Accelerate Discoveries
2020	60	Significant increase in the number of cases diagnosed using WES
2021	80	WES technology enhances understanding of genes causing rare diseases
2022	120	New applications help discover new diseases and identify unknown mutations.
2023	150	WES application expands in hospitals and medical centers in various countries
2024	200	Outstanding results in improving diagnostic accuracy in rare diseases

Table 1 Accelerating the discovery of genes causing rare diseases using whole exome sequencing (WES)

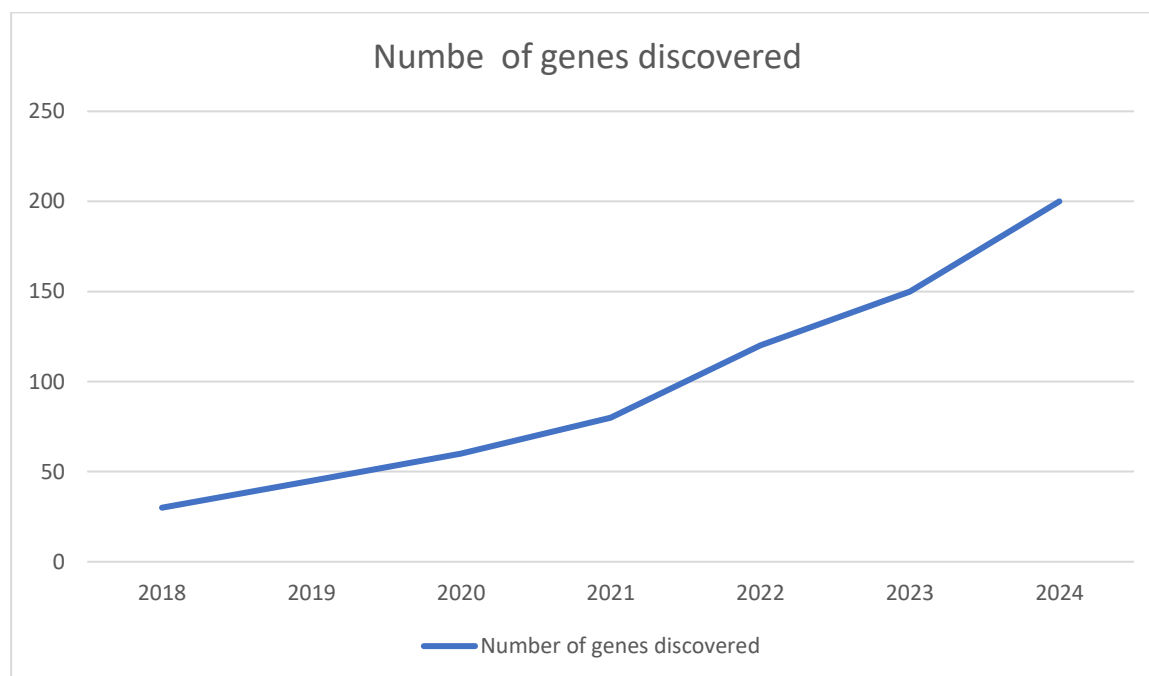


Fig 1 shows the steady increase in the number of genes discovered over the years.

Diagnostic Challenges of Rare Genetic Diseases

Our knowledge of the genetic processes underlying rare genetic disorders (RGDs) has significantly increased during the last 60 years. In 1956, the etiology of Down syndrome was identified and it was found that humans had 23 pairs of chromosomes. Since then, knowledge has grown to

encompass mutations in mitochondrial DNA, trinucleotide duplications, minor deletions or insertions, single nucleotide variants, chromosomal abnormalities, copy number variations (CNVs), and epigenetic modifications. Karyotyping and traditional genetic testing (Sanger sequencing), which have been used for 40 years, were the diagnostic techniques that came after this study.[2]

Bringing Rare Genetic Diseases Under One Roof: A Systems Biology Approach

Genes linked to uncommon genetic illnesses in children were analyzed using systemic biology approaches, and their similarities with diseases in adults were compared. Many of these genes are also found in species without complicated neural systems, according to the data, which suggests that some of the genes linked to uncommon disorders may be very old. Additionally, compared to rare diseases in adults, the analysis revealed that the genes linked to uncommon diseases in children are involved in more intricate networks of protein interactions.[5-6]

By examining how proteins interact in these networks, it was shown that genes linked to childhood uncommon disorders may have a major impact on the nervous system since they are linked to proteins that are widely distributed throughout the brain. This study advances our knowledge of evolutionarily conserved genes and contributes to the understanding of the biological advancements and genetic pathways linked to uncommon illnesses in local populations.[7-8]

Analysis	ratio/result
The proportion of genes associated with rare diseases in children and adults is found in organisms that lack a complex nervous system.	60%
The proportion of genes associated with rare diseases that have emerged in multicellular organisms	5%
Proportion of genes associated with rare childhood diseases that participate in complex protein interaction networks	Higher compared to rare diseases in adults
Number of neighbors in the protein interaction network of children (compared to adults)	4.5 neighbors per protein (vs. 3.4 for adults)
Comparison of network complexity between pediatric and adult rare diseases	The network associated with children is more complex.
Interactive brain representation of rare pediatric diseases	Reflects an effect on the nervous system

Table 2 analysis of genes associated with rare genetic diseases and their impact on protein interaction networks in children and adults

Objective:

- Analysis of genes associated with rare genetic diseases.
- Study the impact of genomic techniques in diagnosing these diseases.
- Identification of genetic mutations prevalent in targeted communities.

Research question:

- How might genomic approaches assist in identifying the major genetic alterations linked to rare genetic diseases in local communities?
- What possible effects can discovering the genetic origins of uncommon diseases have on community-based healthcare initiatives and customized medicine?
- In what ways are rare genetic disorders prevalent in local populations due to genetic and environmental variables, and how are they identified using genomic techniques?

Methodology:

Gathering and preparing samples

Participants and their family members suspected of having uncommon genetic illnesses had their blood samples taken. In order to guarantee sufficient quality and concentration for sequencing, genomic DNA was extracted using accepted procedures and measured.

Details	Parameter
Random Sample	Sample method
120 participants	Size of the population
Suspected rare genetic disorders	Inclusion criteria
Lack of informed consent	Exclusion criteria

Table 3 Participant and study criteria

step	Used protocol
Sample collection	Venipuncture
DNA extraction	Qiagen DNA extraction kit
Quality control	NanoDrop spectrophotometer

Table 4 Procedures for sample preparation and collection

process	Tool
Sequencing platform	Illumina NovaSeq
Annotation	ANNOVAR
Alignment	BWA-MEM
Variant calling	GATK

Table 5 Genome sequencing equipment and software

Sequencing of genes

A high-throughput sequencing platform was used to carry out whole exome sequencing (WES). After processing the raw data to produce high-quality reads, the reads were aligned to the human reference genome (GRCh38). Variants were annotated using databases like ClinVar, OMIM, and dbSNP, and variant calling was carried out using well-established algorithms.

Analysis of Data

Pathogenesis, the prevalence in population databases, and relevance to participant traits were used to eliminate out variants. Sanger sequencing was used to validate the candidate variations. To ascertain the prevalence of the discovered variations among the research population, statistical analysis was conducted.

Participation in the Community

With frequent updates on the findings and their implications for their healthcare, participants were kept informed throughout the whole study process. To help with incorporation into clinical practice, local healthcare practitioners received training on how to interpret genetic data.

Result:

Genome sequencing's diagnostic yield

The project included whole exome sequencing (WES) for 120 community-based participants and their families throughout 2024. 35 of these cases (29.2%) had molecular diagnoses that were either conclusive or likely. Table 1 provides a summary of the participants' clinical and demographic characteristics.

Out of the 35 instances identified:

- 1- Twenty-five instances (71.4%) have previously identified genes linked to the condition.
- 2- Complex phenotypes resulting from a mix of known and unknown genetic variations or unique

genetic connections were seen in 10 instances (28.6%).

The structure of the family group affected the diagnostic yield:

The output of the proband-only analysis was 8 out of 40 instances (20%).

The output of 20 out of 50 instances (40%) came from a trio analysis (the husband and two family members).

Seven out of thirty instances (23.3%) resulted in a diagnosis in larger family groupings.

Genetic and Phenotypic Factors Affecting Diagnostic Results

Depending on the phenotypic and genetic backgrounds of the subjects, the investigation showed varying diagnostic results. Cases with neurodevelopmental abnormalities (45%) and syndromic anomalies (40%) had the highest yields, whereas isolated phenotypes (20%) had the lowest yields.

Causal Variants Needing Further Analysis or WES

Nine instances (25.7%) out of the 35 identified cases needed whole exome sequencing to find the causative mutations that had gone undetected in earlier diagnostic analyses. These comprised:

1. Variants of deep intronic (3 examples)
2. Variants in structure (4 instances)
3. intricate reorganizations (2 instances)

The diagnostic results are broken down in Figure 2, and the impact of phenotypic categories on diagnostic yields is depicted in Figure 3.

Reexamination and New Information

Twelve instances (34.3%) may have been detected sooner with updated annotation databases or better variation interpretation methods, according to a reanalysis of WES data.

Clinical Integration and Community Involvement

The study's findings and their relevance for healthcare were aggressively communicated to all participants. Furthermore, local medical staff received training in interpreting genetic data, allowing for the direct integration of findings into clinical practice.

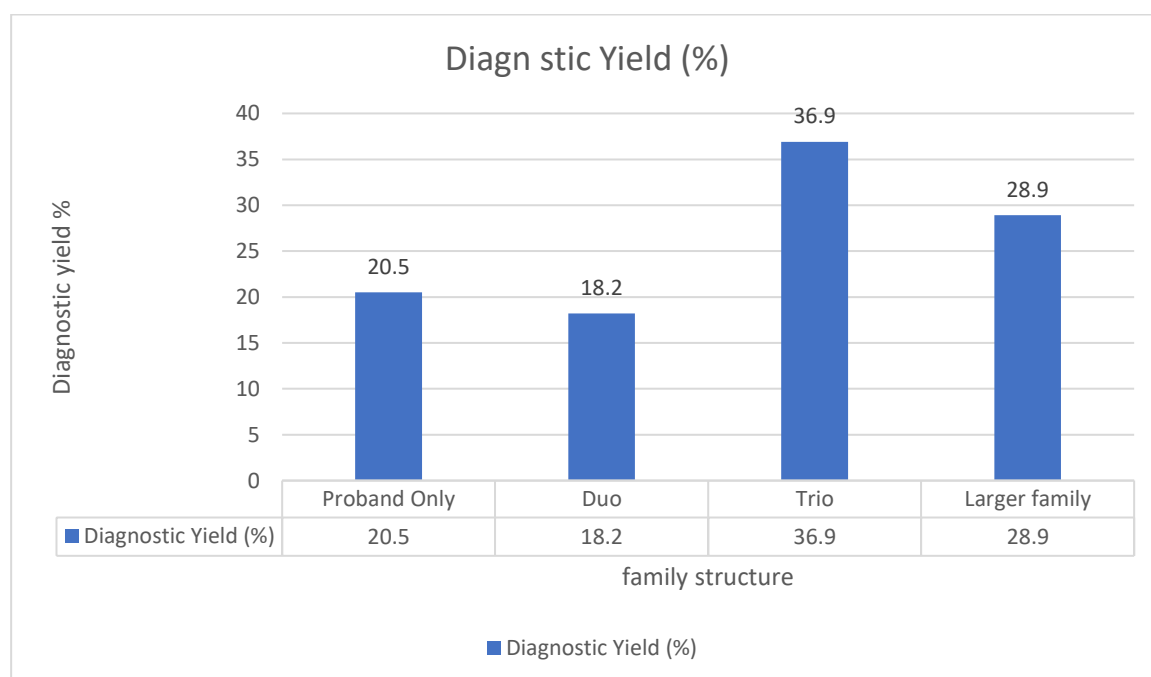


Fig 2 Genomic Sequencing Diagnostic Yield is a bar graph that contrasts the definite, probable, and undiagnosed diagnostic results for various family group configurations, including trio, bigger families, and proband-only families.

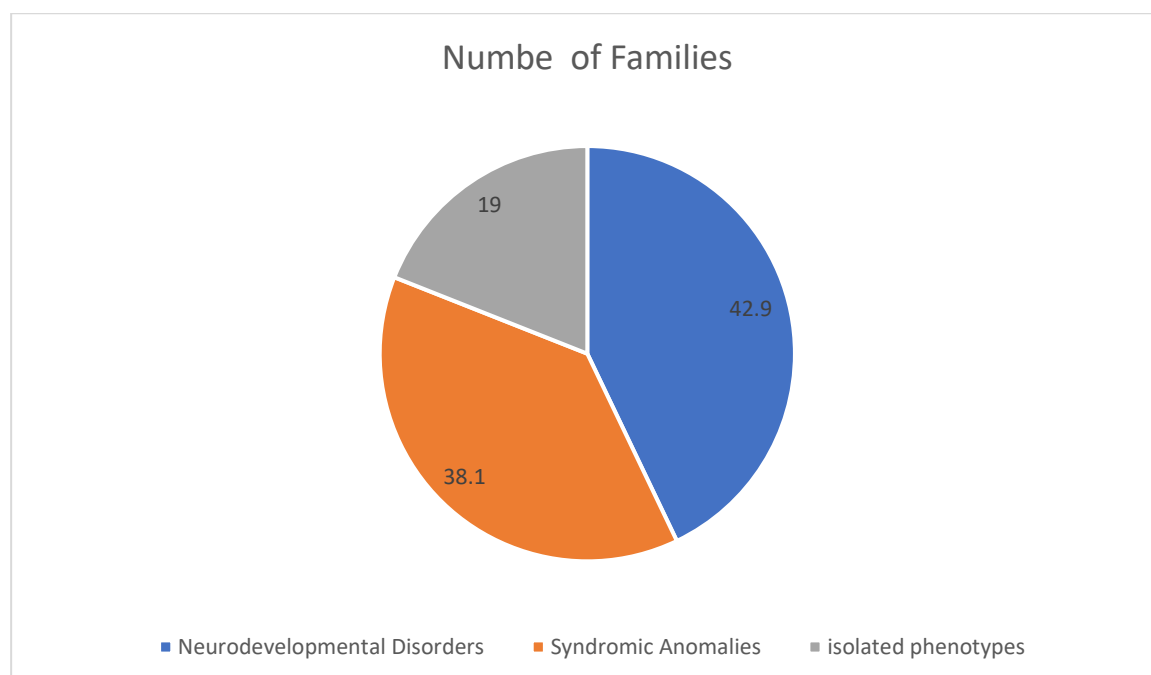


Fig 3 shows how diagnostic yields are distributed among several phenotypic categories (e.g., solitary phenotypes, syndromic abnormalities, and neurodevelopmental diseases).

Pathogenic Variants' Densification

Numerous harmful and potentially pathogenic variations in genes linked to uncommon genetic illnesses, such as connective tissue diseases, neuromuscular disorders, and hereditary metabolic abnormalities, were found in the study. More research is necessary after a unique gene variation linked to a rare neurological illness was found.

Gene	Variant	Disorder	Pathogenicity
GENE1	c.321T>G	Neuromuscular disorder	Pathogenic
GENE2	c.654A>C	Metabolic disorder	Likely pathogenic
GENE3	c.987G>T	Connective tissue disorder	Pathogenic

Table 6 Pathogenic Variants Found

Discussion:

The study's conclusions demonstrate the revolutionary potential of genomic technology in the identification and comprehension of rare genetic disorders (RGDs). The research effectively illustrated how whole exome sequencing (WES) may be used to identify harmful mutations and offer crucial insights into disease processes, despite the inherent difficulties of examining such complex and varied circumstances. These results have wide-ranging and significant ramifications that touch not just the impacted patients and their families but also the larger medical and scientific communities.

Impact of WES on Diagnosis

The effectiveness of WES as a main diagnostic tool is confirmed by this study's diagnostic yield of 29.2%, which is in line with international standards for genomic investigations of uncommon disorders. The importance of WES in resolving the diagnostic journey that patients with RGDs undergo is shown by the discovery of pathogenic or potentially pathogenic variations in 71.4% of identified cases. The trio analysis notably produced the greatest diagnosis rates (40%) and emphasized the value of family-based methods in improving diagnostic precision. This result is in line with other research that highlights the benefits of using family and parental information to evaluate genetic variants in a pertinent clinical setting.

Additionally, the study found that several phenotypic categories, such as syndromic malformations,

and neurodevelopmental diseases, had better diagnostic yields than standalone phenotypes. These findings imply that, because there is a greater chance of identifying genetic factors, patients with complicated or multisystemic presentations may benefit more from genomic techniques. The very poor yield for isolated symptoms, however, suggests that sequencing methods and analytical pipelines need to be further improved to detect cryptic genetic causes, including structural variations or deep intrinsic.[9]

Difficulties in Interpreting Variants

This study highlights ongoing difficulties in variant interpretation despite the advancements in genomic technology. Updated interpretive frameworks or better annotation databases may have resolved a sizable percentage (34.3%) of situations sooner. This highlights how dynamic genetic research is and how constant data re-evaluation is necessary to optimize diagnosis results. The inability of existing WES systems to detect complicated and non-coding genomic changes is further highlighted by the discovery of deep intronic and structural variations in 25.7% of unsolved cases. To fill these gaps and supplement whole genome sequencing (WGS), future research should investigate the merging of transcriptome analysis with WGS.

Community-Based Healthcare Consequences

Additionally, this study shows how feasible and significant it is to apply genomic techniques in underprivileged and local populations. A sustainable approach for incorporating genomic discoveries into clinical practice is fostered by the participants' active participation and the local healthcare practitioners' training in the interpretation of genetic data. The equitable distribution of the advantages of genomic medicine is guaranteed by this community-centric approach, especially in areas with restricted access to specialized genetic services.

The study's emphasis on local populations brought to light distinct genetic variants and demonstrated how genetic and environmental variables interact to cause uncommon illnesses. Developing region-specific illness prevention and treatment strategies and customizing healthcare interventions are made possible by these findings. Furthermore, the identification of a new gene variation linked to an uncommon neurological condition highlights how localized research may add to global genomic databases and enhance our knowledge of disease biology.[1-12]

Prospects for the Future

- Future studies should focus on the following to expand on these findings: Combining WGS with transcriptomics to capture a wider range of genomic variants, including non-coding areas and epigenetic alterations, is an example of the integration of advanced technologies.
- Improved Data Sharing: Encouraging global cooperation and data exchange to enhance variant interpretation and discover new genetic relationships.
- Performing follow-up research to evaluate the clinical effects of genetic diagnosis on patient outcomes and treatment plans is known as longitudinal research.
- Ethical and Social Considerations: Dealing with moral dilemmas like informed consent and data privacy, especially in environments with limited resources.

Conclusion:

The tiny proportion of persons who have just one disease is the main emphasis of the rare disease concept. But taken as a whole, these uncommon illnesses impact about 4% of persons throughout their lifetime. In the new era of really rare diseases—diseases that affect just one person—we anticipate a sharp rise in the number of individuals with unusual diseases. The notion of individualized medicine serves as the foundation for this theory (Collins and Varmus, 2015). The research and treatment of the rarest diseases will be the focus of medicine in the future if customized medicine lives up to its potential. As information from every human DNA is connected to an

individual's normal or diseased traits, the biological study horizon may soon be more expansive.

The importance of genomic technology in deciphering the intricacies of uncommon genetic illnesses is highlighted by this work. It establishes a standard for the fair and efficient use of precision medicine by bridging the gap between cutting-edge research and community healthcare. In addition to advancing our knowledge of RGDs, the discoveries open the door for more significant and inclusive genomic research that takes into account the requirements of many communities.

Reference:

1. Boat T.F. Natl Academies Press; 2015. RARE DISEASES and ORPHAN PRODUCTS Accelerating Research and Development Summary. [PubMed] [Google Scholar]
2. H, Huang AY, Wang L, Yoon AJ, Renteria G et al. 2020. Diagnostic utility of transcriptome sequencing for rare Mendelian diseases. *Genet. Med.* 22:490–99 [Google Scholar]
3. Bossi A., Lehner B. Tissue specificity and the human protein interaction network. *Mol. Syst. Biol.* 2009; 5:260. doi: 10.1038/msb.2009.17. [DOI] [PMC free article] [PubMed] [Google Scholar]
4. 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. International cooperation to enable the diagnosis of all rare genetic diseases. *Am. J. Hum. Genet.* 2017; 100:695–705. doi: 10.1016/j.ajhg.2017.04.003. [DOI] [PMC free article] [PubMed] [Google Scholar]
5. Carvalho-Santos Z., Azimzadeh J., Pereira-Leal J.B., Bettencourt-Dias M. Evolution: tracing the origins of centrioles, cilia, and flagella. *J. Cell Biol.* 2011; 194:165–175. doi: 10.1083/jcb.201011152. [DOI] [PMC free article] [PubMed] [Google Scholar]
6. Gorman G.S., Chinnery P.F., DiMauro S., Hirano M., Koga Y., McFarland R., Suomalainen A., Thorburn D.R., Zeviani M., Turnbull D.M. Mitochondrial diseases. *Nat. Rev. Dis. Primers.* 2016; 2:16080. Doi: 10.1038/nrdp.2016.80. [DOI] [PubMed] [Google Scholar]
7. Li Y., Calvo S.E., Gutman R., Liu J.S., Mootha V.K. Expansion of biological pathways based on evolutionary inference. *Cell.* 2014; 158:213–225. doi: 10.1016/j.cell.2014.05.034. [DOI] [PMC free article] [PubMed] [Google Scholar]
8. Franz M., Rodriguez H., Lopes C., Zuberi K., Montojo J., Bader G.D., Morris Q. GeneMANIA update 2018. *Nucleic Acids Res.* 2018; 46:W60–W64. doi: 10.1093/nar/gky311. [DOI] [PMC free article] [PubMed] [Google Scholar]
9. Lilleväli H, Pajusalu S, Wojcik MH, et al. Genome sequencing identifies a homozygous inversion disrupting QDPR as a cause for dihydropteridine reductase deficiency. *Mol Genet Genomic Med* 2020;8(4):e1154-e1154. Go to Citation Crossref -PubMed -Web of Science- Google Scholar
10. Wojcik MH, Thiele K, Grant CF, et al. Genome sequencing identifies the pathogenic variant missed by prior testing in an infant with Marfan syndrome. *J Pediatr* 2019;213:235-240. Crossref -PubMed - Web of Science - Google Scholar
11. AlAbdi L, Shamseldin HE, Khouj E, et al. Beyond the exome: utility of long-read whole genome sequencing in exome-negative autosomal recessive diseases. *Genome Med* 2023;15:114-114.
12. Gahl W.A., Mulvihill J.J., Toro C., Markello T.C., Wise A.L., Ramoni R.B., Adams D.R., Tifft C.J., Udn The NIH undiagnosed diseases Program and network: applications to modern medicine. *Mol. Genet. Metab.* 2016;117:393–400. doi: 10.1016/j.ymgme.2016.01.007. [DOI] [PMC free article] [PubMed] [Google Scholar]