# The clinical utility of genome sequencing in the molecular diagnosis of genes related to inborn errors of metabolism

Xiangwen Chen-Deutsch, Mary Colasanto, Cristina da Silva, Christin Collins, Madhuri Hegde

## 1 Introduction

In the era of clinical and molecular genetics and genomics, disorders of inborn errors of metabolism (IEM) can no longer be considered limited to abnormalities in the synthesis or catabolism of molecules in pathways measurable in classical biochemical assays. Instead, they are defined as impairments of biochemical pathways intrinsically resulting in the pathophysiology of various diseases. While IEM disorders are routinely diagnosed via biochemical testing, individuals with early onset of nonspecific phenotypes, unusual presentations, or multiple co-occurring genetic disorders may undergo extensive testing before receiving a diagnosis. The clinical utility of genome sequencing is quickly being realized for various indications, and this study focuses on a thorough investigation in the context of IEM diseases.

## 2 Methods

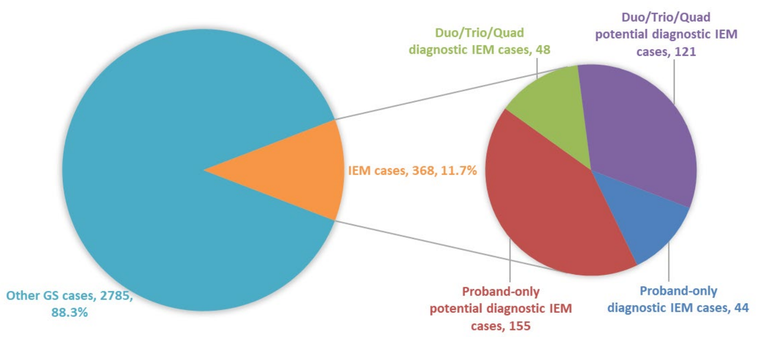
Genome sequencing (GS) was performed on genomic DNA using 2X150bp reads by next-generation sequencing (NGS) at a mean coverage of 40X across the entire genome. A retrospective investigation of reported IEM-related genetic variants was conducted in a total of 3,153 GS cases, with provided information on phenotypes, from a continuous period. The age of the probands ranged from newborn to 80 years.

## 3 Results

Of the 3,153 GS cases, 368 (11.7%) reported IEM-related gene variants with diagnostic or potential diagnostic findings (Figure 1). Diagnostic cases are defined as those with pathogenic/likely pathogenic (P/LP) variants reported, while potential diagnostic cases are defined as those with variants of uncertain significance (VUS) and/or P/LP variants reported. This proportion aligns with previous estimates that IEMs account for a significant portion of genetic disorders, typically ranging from 10-15% in most studies.

Among the 368 reported IEM cases, 199 (54.1%) underwent proband-only GS testing, while 169 (45.9%) opted for family-based testing (Trio, Duo, or Quad) (Figure 1). This distribution differs from some earlier studies that predominantly used trio-based analyses, indicating a shift towards more diverse testing strategies.

*Figure 1: Numbers and percentages of reported IEM cases out of 3,153 total GS cases.*




Of the 1,021 IEM-related genes investigated in this GS cohort, 214 (20.9%) were reported in these cases (Table 1), indicating a substantial genetic diversity within IEM disorders and highlighting the importance of comprehensive genetic testing for potential IEM diagnosis.

*Table 1: Number of IEM-related genes reported in different categories of IEM disorders.*

| Category | Number of genes reported |
|---|---|
| Disorders of nitrogen-containing compounds (e.g. amino acids and nucleotide metabolism disorders) | 42 |
| Mitochondrial disorders of energy metabolism (nuclear related genes) | 38 |
| Disorders of lipids (e.g. carnitine, lipoprotein, and sterol and steroid metabolism disorders) | 28 |
| Disorders of vitamins, cofactors, metals and minerals | 27 |
| Congenital disorders of glycosylation | 26 |
| Storage disorders (e.g. lysosomal storage disease, and autophagy disorders) | 21 |
| Disorders of carbohydrates (e.g. glycogen storage diseases, and insulin secretion and signaling disorders) | 18 |
| Disorders of tetrapyrroles (e.g. heme and bilirubin metabolism disorders) | 8 |
| Disorders of peroxisomes and oxalate | 6 |
| Total | 214 |

*Table 2: Frequently reported IEM-related genes (>5 cases), their associated disorders, and metabolic pathways.*

| IEM Pathway | Gene | Enzyme or protein | Inheritance | OMIM Phenotype | No. of reported cases |
|---|---|---|---|---|---|
| Disorder of the pentose phosphate pathway and polyol metabolism | *G6PD* | Glucose-6-phosphate dehydrogenase | XL | Anemia, congenital, nonspherocytic hemolytic, 1, G6PD deficient | 7 |
| Disorder of glutamate metabolism | *GRIN2B* | Glutamate receptor, ionotropic, N-methyl-D-aspartate (NMDA), subunit 2B | AD | Intellectual developmental disorder, autosomal dominant 6, with or without seizures (AD); Developmental and epileptic encephalopathy 27 (AD) | 7 |
| Disorder of β- and γ-amino acids | *GABBR2* | Gamma-aminobutyric acid (GABA) B receptor 2 | AD | Neurodevelopmental disorder with poor language and loss of hand skills (AD); Developmental and epileptic encephalopathy 59 (AD) | 6 |
| Disorder of glutamate metabolism | *GRIN2A* | Glutamate receptor, ionotropic, N-methyl-D-aspartate (NMDA), subunit 2A | AD | Epilepsy, focal, with speech disorder and with or without impaired intellectual development | 6 |
| Disorder of glutamate metabolism | *GRIN2D* | Glutamate receptor, ionotropic, N-methyl-D-aspartate (NMDA), subunit 2D | AD | Developmental and epileptic encephalopathy 46 | 6 |
| Disorder of nucleotide metabolism | *IFIH1* | Interferon-induced helicase C domain-containing protein 1 | AD, AR | Singleton-Merten syndrome 1 (AD); Aicardi-Goutieres syndrome 7 (AD); Immunodeficiency 95 (AR) | 6 |
| Disorder of lipoprotein metabolism | *LDLR* | Low density lipoprotein receptor | AD, AR | Hypercholesterolemia, familial, 1 | 6 |
| Disorder of mitochondrial DNA depletion or intergenomic communication | *POLG* | Polymerase, DNA, gamma | AD, AR | Progressive external ophthalmoplegia, autosomal dominant 1, and autosomal recessive 1; Mitochondrial DNA depletion syndrome 4A (Alpers type), and 4B (MNGIE type) (AR); Mitochondrial recessive ataxia syndrome (includes SANDO and SCAE) (AR) | 6 |
| Disorder of lipoprotein metabolism | *APOE* | Apolipoprotein E | AD, AR | Alzheimer disease 2 (AD); Sea-blue histiocyte disease (AR); Lipoprotein glomerulopathy (AD); Hyperlipoproteinemia, type III (AR, AD) | 5 |
| Disorder of carnitine metabolism | *CPT2* | Carnitine palmitoyltransferase II | AR, AD | CPT II deficiency, myopathic stress-induced (AD, AR), infantile (AR), and lethal neonatal (AR) | 5 |
| Disorder of cobalamin metabolism | *HCFC1* | Host cell factor C1 | XLR | Methylmalonic aciduria and homocysteinemia, cblX type | 5 |
| Disorder of phosphoinositide metabolism | *ITPR1* | Inositol 1,4,5-triphosphate receptor, type 1 | AD, AR | Spinocerebellar ataxia 15, and 29 congenital nonprogressive (AD); Gillespie syndrome (AD, AR) | 5 |
| Glycosylation disorder of vesicular trafficking | *VPS13B* | Vacuolar protein sorting 13 homolog B | AR | Cohen syndrome | 5 |

Among the reported IEM-related gene variants, AR cases had 97 P/LP and 162 VUS, AD cases had 14 P/LP and 68 VUS, AR/AD or SD cases had 48 P/LP and 80 VUS, and XL cases had 17 P/LP and 33 VUS (Figure 3). Notably, AR cases showed the highest number of both P/LP and VUS variants, consistent with the predominance of AR inheritance in IEM disorders.

Thirteen IEM-related genes were reported in more than five cases (Table 2), highlighting their potential diagnostic importance. While *G6PD, CPT2,* and *POLG* are well-established in IEM literature, others like *IFIH1* are less commonly reported. Notably, the list includes genes typically associated with neurodevelopmental disorders (e.g., *GABBR2, GRIN2A/B/D*), suggesting an expanding definition of IEM disorders.

Among diagnostic/potential diagnostic cases, 39 UTR and deep intronic variants were reported (3 classified as P/LP, 5 in 5'UTR, 5 in 3'UTR, 29 in deep intronic regions), typically undetectable by exome-based sequencing. These non-coding variants are crucial for identifying secondary disease-causing variants in autosomal recessive IEM disorders. Despite challenges in determining pathogenicity due to limited databases and complex functional studies not routinely performed in clinical settings, genome sequencing's ability to capture regulatory elements potentially increases the diagnostic yield of IEM disorders.

Of the 92 diagnostic cases, 45.7% were autosomal recessive (AR), 15.2% autosomal dominant (AD), 29.3% AR/AD or semidominant (SD), and 9.8% X-linked (XL). For the 276 potential diagnostic cases, the distribution was 34.4% AR, 24.6% AD, 28.6% AR/AD or SD, and 12.3% XL (Figure 2). The differing distribution between diagnostic and potential diagnostic cases may reflect varying certainty in variant interpretation across inheritance patterns. Notably, AR is the predominant inheritance pattern in IEM disorders, previously reported in approximately 70% of known IEM disorders.

*Figure 2: Inheritance patterns in diagnostic and potential diagnostic IEM cases.*

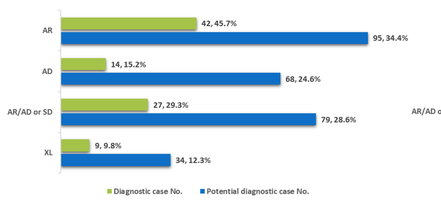
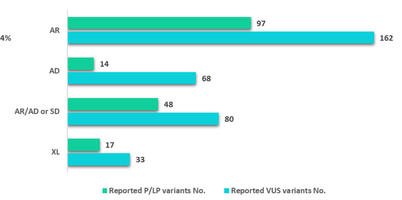



*Figure 3: Numbers of reported IEM-related P/LP and VUS variants in cases with different inheritance patterns.*




## 4 Conclusions

A wide spectrum of clinical symptoms in patients often leads to enigmas in clinical diagnosis, making genome-level molecular diagnosis a critical step in assisting the clinical world. The study of this large cohort of GS cases revealed that IEM disorders represent an important proportion of diseases that can be readily diagnosed at the molecular level by GS. GS testing for the proband with family members reduced the molecular diagnosis burden during the data analysis process compared to proband-only testing, especially for recessive conditions, which is the most common inheritance pattern for IEM disorders. Furthermore, disease-causing genetic variations can be identified more efficiently by GS compared to other large-scale genomic assays such as exome sequencing. In follow-up, the profile of molecular diagnosis of IEM disorders also helps the clinic gain a deeper understanding of the mechanisms and manifestations of these genetic diseases.

*Reference*

Ferreira CR, van Karnebeek CDM, Vockley J, Blau N. A proposed nosology of inborn errors of metabolism. Genet Med. 2019 Jan;21(1):102-106. doi: 10.1038/s41436-018-0022-8. Epub 2018 Jun 8. PMID: 29884839; PMCID: PMC6286709.