

# Impacts of New Mutations involved in Metabolic processes Leading to Intellectual Disabilities Found by Whole-Exome Sequencing and *In Silico* Investigations

Vahid Omarmeli<sup>1</sup>, PhD candidate; Elham Siasi<sup>2</sup>, PhD; Maryam Peyvandi<sup>3</sup>, PhD; Sahar Mansoubi<sup>4</sup>, PhD; Mohaddeseh Mohsenpour<sup>2</sup>, PhD

<sup>1</sup>Department of Genetics, NT.C., Islamic Azad University, Tehran, Iran;

<sup>2</sup>Department of Microbiology, NT.C., Islamic Azad University, Tehran, Iran;

<sup>3</sup>Department of Biology, NT.C., Islamic Azad University, Tehran, Iran;

<sup>4</sup>Cellular and Molecular Biology, NT.C., Islamic Azad University, Tehran, Iran

## Correspondence:

Elham Siasi, PhD;  
Vafadar Blvd., Shahid Sadoughi St.,  
Hakimiyeh Exit, Shahid Babaee Highway,  
Postal code: 1651153311, Tehran, Iran  
Tel: +98 21 77009801-42

Email: el\_siasi@iau.ac.ir

Received: 19 November 2025

Revised: 28 May 2026

Accepted: 24 June 2026

## What's Known

- The genetic etiology of Intellectual disabilities (IDs) has been reported to account for about two-thirds, comprising chromosomal abnormalities, monogenic diseases, polygenic diseases, epigenetic anomalies, hereditary metabolic conditions, etc.
- At present, detected hereditary metabolic conditions represent 1-5% of ID etiology. More genetic and metabolic diseases are becoming identifiable and curable thanks to advancements in gene analysis and diagnostic technologies.

## What's New

- This study identified 73 novel mutations among 139 mutations detected across 56 cases. Of these cases, 31 had a history of consanguinity.
- All novel mutations were related to 69 genes and a candidate group (19 genes), which were selected for further analyses. hsa-miR-615-3p was found as the top miRNA.

## Abstract

**Background:** Inborn errors of metabolism (IEM), a notable subset of intellectual disabilities (IDs), may increase the likelihood of new mutation hotspots. Thus, this study aimed to investigate whole-exome sequencing (WES) reports of complicated cases.

**Methods:** The *ATP7B* mutation (Gly205Glu) was selected for homology modeling with MODELLER, then superimposed with Matchmaker. Moreover, an *in silico* strategy was designed to identify new molecular relationships among the novel mutations using STRING-MODEL, NetworkAnalyst, Enrichr, and Metascape.

**Results:** A total of 114 cases that underwent WES testing were included (2023-2025, Kermanshah, Iran). Results showed 73 novel mutations (out of 139) in 56 cases, of which 31 cases had consanguineous backgrounds. All novel mutations were associated with 69 genes, and a candidate group of 19 genes was selected for further analyses. Gene-miRNA interactions predicted hsa-miR-615-3p as the top miRNA. Enrichment Analysis indicated potential processes, including Alanine, aspartate, and glutamate metabolism ( $P=5.6E-4$ ); Mitochondrial Matrix ( $P=3.06E-4$ ); Mental Retardation ( $P=3.11E-10$ ); and ID ( $P=2.34E-08$ ). Notably, the Alpha-amino acid metabolic process was the best-scored process by Clustering Enrichment Analysis.

**Conclusion:** Altogether, this study reported novel mutations, identified a candidate gene list, and suggested that metabolic processes may play pivotal roles in disease causation through new mutations leading to ID manifestations.

Please cite this article as: Omarmeli V, Siasi E, Peyvandi M, Mansoubi S, Mohsenpour M. Impacts of New Mutations involved in Metabolic processes Leading to Intellectual Disabilities Found by Whole-Exome Sequencing and *In Silico* Investigations. Iran J Med Sci. doi:

**Keywords** • Intellectual disability • Metabolic disorders • Exome sequencing • Novel mutation • Systems biology

## Introduction

Pediatricians are frequently referred to treat patients with global developmental delay (GDD) and intellectual disabilities (IDs), which impact 2% to 3% of children.<sup>1</sup> Several published guidelines recommend different approaches for evaluating the etiology of GDD/ID.<sup>1-3</sup> A prompt and precise diagnosis may terminate the diagnostic odyssey, reveal a curable cause, and give prognostic data. It can also improve accessibility to community resources.

To enable faster treatment and better health outcomes, preliminary testing might additionally detect an inborn error of metabolism (IEM) before the phenotype progresses to more severe neurologic and/or systemic signs.<sup>4</sup>

IDs have a complicated etiology that can be categorized into two distinct groups: genetic and non-genetic.<sup>5</sup> The genetic etiology accounts for about two-thirds, comprising chromosomal abnormalities, monogenic diseases, polygenic diseases, epigenetic anomalies, hereditary metabolic conditions, and so on.<sup>6</sup> Although genetic risk factors may be more prevalent among individuals with GDD/ID, they alone cannot fully explain its complex pathogenesis or the considerable phenotypic heterogeneity observed among affected individuals.<sup>3</sup>

Currently, diagnosed hereditary metabolic conditions are estimated to account for approximately 1-5% of the etiologies of ID.<sup>7</sup> More genetic and metabolic diseases are becoming identifiable and curable thanks to advancements in gene analysis and diagnostic technologies, which brings hope for treating the etiology of pediatric GDD/ID.<sup>8</sup> A thorough assessment of the literature by van Karnebeek and others identified 81 curable IEMs with ID.<sup>9</sup> The potential for improved health outcomes through timely therapeutic intervention has led to the development of tiered diagnostic approaches that prioritize identifying curable diseases.<sup>7</sup>

Whole-exome sequencing (WES) is now clinically available for detecting GDD/ID, with diagnostic yields ranging from 30% to 55%. Recent evidence regarding its clinical applications, including prognostic data, time to diagnosis, pharmacological treatment, and family planning,

further supports its increasing clinical accessibility. Consequently, referral to a clinical geneticist and/or a metabolic specialist may allow further targeted metabolic and genetic testing based on specific clinical characteristics, leading to additional diagnoses, especially for children with ID/GDD and neurological features in whom no etiology is identified through initial high-level investigations.<sup>10,11</sup> During a two-year clinical experience evaluating patients referred to our medical genetics lab for WES testing, our team developed a hypothesis based on the relationships between ID phenotypes and metabolic manifestations among patients with novel mutations and severe, potentially life-threatening conditions. To test the hypothesis, we designed this research based on systematic scientific data collection, data filtering, and categorization, and, finally, *in silico* analysis of novel mutations.

## Material and Methods

### Sampling and Data Collection

The inclusion criteria for performing a WES test in suspected cases with potential disease-causing genetic mutations were as follows: initial clinical suspicion by a physician based on clinical examination; recommendation by a genetic consultant; family history; biochemical examination results; maternal history of abortion; presence of affected siblings; and severity of clinical manifestations, including facial and skeletal features, speech, and communication skills. Clinical characteristics were also categorized according to age, sex, consanguinity status, zygosity, and inheritance patterns (table 1).

**Table 1:** Clinical characteristics of all cases included in the current study based on their initial information

| Clinical characteristics         | Category | Duration/Number |
|----------------------------------|----------|-----------------|
| Age (n=114)                      | Min      | 6 months        |
|                                  | Max      | 58Y             |
|                                  | mean±SD  | 18.82±13.15Y    |
| Sex (n=114)                      | Male     | 56 (49.12%)     |
|                                  | Female   | 58 (50.88%)     |
| Consanguinity (n=114)            | Yes      | 64 (56.14%)     |
|                                  | No       | 50 (43.86%)     |
| Zygosity (n=139 mutations)       | Hom      | 57 (41%)        |
|                                  | Het      | 79 (56.84%)     |
|                                  | Hemi     | 3 (2.16%)       |
| Inheritance (n=120 unique genes) | AD       | 25 (20.83%)     |
|                                  | AR       | 83 (69.17%)     |
|                                  | AD/AR    | 9 (7.5%)        |
|                                  | XLD      | 2 (1.66%)       |
|                                  | XLR      | 1(0.84%)        |
| Novel mutation (n=139 mutations) | Yes      | 73 (52.52%)     |
|                                  | No       | 66 (47.48%)     |

Data on age, sex, and consanguinity status were available for all 114 cases. The total number of zygosity and novel mutations was 142, corresponding to the 142 detected mutations. The total number of inheritance records was 120, corresponding to the 120 unique genes identified among all subjects. Min: Minimum; Max: Maximum; Ave: Average; Hom: Homozygote; Het: Heterozygote; Hem: Hemizygote; AD: Autosomal dominant; AR: Autosomal recessive; XLD: X-linked dominant; XLR: X-linked recessive

Written informed consent was obtained from all cases (older than 18 years) or their parents to participate in the study, with personal data kept anonymous. This research was consistent with the Declaration of Helsinki and approved by the Research Committees of Islamic Azad University-North Tehran Branch with a certified ethical number (IR.IAU.TNB.REC.1404.033).

#### Whole-Exome Sequencing (WES) Analysis

To detect disease-causing mutations in the cases, the WES test was performed step by step as follows: Genomic deoxyribonucleic acid (gDNA) was extracted and purified from all individuals' specimens using a DNA extraction kit (Blood Genomic DNA Extraction mini-Kit, FAVORGEN) according to the manufacturer's protocol. All extracted samples were then quantified. For DNA preparation, a total of 1.0 µg gDNA was used per sample. Sequencing libraries were prepared using the Agilent SureSelect Human All ExonV7 Kit (Agilent Technologies, CA, USA) according to the manufacturer's instructions; then, x-index codes were added to the sample attribute sequences. In summary, the fragmentation process was accomplished by the hydrodynamic shear system (Covaris, Massachusetts, USA), producing 180-280 bp fragments. Remaining overhangs were blunted by exonuclease/polymerase activities, and enzymes were removed. After adenylation of the 3' ends of the DNA fragments, adapter oligonucleotides were ligated. DNA fragments containing adapter molecules at both ends were subsequently selectively enriched by polymerase chain reaction (PCR) amplification. Captured libraries were further enriched through a PCR reaction to add index tags for hybridization preparation. Using the AMPure XP system (Beckman Coulter, Beverly, USA), the products were size-selected and quantified by the Agilent High Sensitivity DNA Assay on the Agilent Bioanalyzer 2100 System. Quality-passed libraries were loaded onto the NovaSeq 6000 Illumina sequencers.

Data quality control, analysis, and interpretation were then performed on Generation G9 (the HP server) applying a Unix-based operating system. Notably, the depth of coverage certified for this WES testing was 200x.

#### In Silico Analyses

A two-step *in silico* investigation was conducted on novel mutations refined from the WES results of 114 cases. The first step was carried out by STRING-MODEL (<https://string-db.org/>) for protein-protein interactions (PPIs) and NetworkAnalyst (<https://www.networkanalyst.ca/NetworkAnalyst/>) for gene regulatory networks (GRNs). Notably, GRNs were divided into three categories, including gene-miRNA interactions (GMIs), transcription factor-gene interactions (TFGIs), and protein-chemical interactions (PCIs). The second step consisted of two main assessments: Enrichment analysis (EA) and clustering enrichment analysis (CEA) by Enrichr (<https://maayanlab.cloud/Enrichr/>) and Metascape (<https://metascape.org/gp/index.html#/main/step1>), respectively. All details and related references used in the *in silico* tools are summarized in table 2.

#### Statistical Analysis

Using Enrichr (<https://maayanlab.cloud/Enrichr/>), CEA was performed as the second step of *in silico* analysis on the candidate gene list across three categories: pathway analysis, gene ontologies (GO), and diseases/drugs analysis. For the gene of interest, EA was performed using the following ontology and pathway databases: KEGG Pathway, Reactome Gene Sets, GO Biological Processes, Canonical Pathways, WikiPathways, and CORUM, with all genes used as the enrichment background. Terms with a P value <0.01, a minimum count of 3, and an enrichment factor >1.5 were gathered and grouped into clusters according to their membership similarities. The enrichment factor refers to the ratio between the observed counts and the counts expected by chance.

**Table 2:** Tools, validations, and references of both main *in silico* databases used in this study

| Level | Database             | Site                                                                                                                                                            | Software (version)   |
|-------|----------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------|
| PPIs  | STRING-MODEL         | <a href="https://string-db.org/">https://string-db.org/</a>                                                                                                     | STRING (12.0)        |
| GRNs  | GMIs miRTarBase      | <a href="https://mirtarbase.cuhk.edu.cn/~miRTarBase/miRTarBase_2022/php/index.php">https://mirtarbase.cuhk.edu.cn/~miRTarBase/miRTarBase_2022/php/index.php</a> | NetworkAnalyst (3.0) |
|       | TF-Gene Interactions | <a href="https://www.encodeproject.org/">https://www.encodeproject.org/</a>                                                                                     | NetworkAnalyst (3.0) |
|       | PCIs                 | <a href="https://ctdbase.org/">https://ctdbase.org/</a>                                                                                                         | NetworkAnalyst (3.0) |
| EA    | Pathway Analysis     | <a href="https://maayanlab.cloud/Enrichr/">https://maayanlab.cloud/Enrichr/</a>                                                                                 | Enrichr              |
|       | GO                   | <a href="https://maayanlab.cloud/Enrichr/">https://maayanlab.cloud/Enrichr/</a>                                                                                 | Enrichr              |
|       | DDA                  | <a href="https://maayanlab.cloud/Enrichr/">https://maayanlab.cloud/Enrichr/</a>                                                                                 | Enrichr              |
| MA    | CEO                  | <a href="https://metascape.org/gp/index.html#/main/step1">https://metascape.org/gp/index.html#/main/step1</a>                                                   | Metascape            |

PPIs: Protein-protein interactions; GRNs: Gene regulatory networks; GMIs: Gene-miRNA interactions; TFGIs: Transcription factor-gene interactions; PCIs: Protein-chemical interactions; EA: Enrichment analysis; GO: Gene ontology; DDA: Disease-drug assessment; VAA: Variant annotation Assessment; MA: Meta-analysis; CEP: Clustering enriched ontology

More precisely, P values were calculated using the cumulative hypergeometric distribution,<sup>12</sup> and Q values were calculated using the Benjamini-Hochberg method to correct for multiple testing.<sup>13</sup> Kappa scores<sup>14</sup> were used as a similarity measure for calculating hierarchical clustering of enriched terms, and sub-trees with a similarity score >0.3 were defined as clusters. The most statistically significant term within each cluster was selected as the representative term.

### *Mutation Confirmation*

As the gold-standard method for validating the WES findings, we employed Sanger sequencing to confirm the genotypes of individuals carrying the identified novel mutation. For each novel mutation, Sanger screening was performed by designing mutation-specific primers, conducting the sequencing procedure, and analyzing the resulting chromatograms using Chromas version 2.6.6. Representative chromatograms of selected novel mutations are presented in the Results section.

### *Homology Modeling and Superimposition Investigations*

To gain clearer insight into the potential pathogenic or likely pathogenic effects of novel mutations identified by NGS analysis, we designed a workflow to assess the functional implications of selected novel mutations identified in this study. The workflow included homology modeling by retrieving the canonical protein sequence of the corresponding gene from UniProtKB (<https://www.uniprot.org/>), generating separate wild-type and mutant protein sequences, constructing three-dimensional protein structures using I-TASSER-MTD (<https://zhanggroup.org/I-TASSER-MTD/>), SWISS-MODEL (<https://swissmodel.expasy.org/>), and MODELLER (version 10.6), and energy-minimizing the resulting structures using Swiss-PdbViewer (ver. 4.1.1). Structure validation was performed using Ramachandran plot analysis in MolProbity (<http://molprobity.biochem.duke.edu/index.php>) and z-score analysis in ProSA (<https://prosa.services.came.sbg.ac.at/prosa.php>). Following quality control, the wild-type and mutant structures were superimposed using the Matchmaker tool in Chimera (version 1.18) to assess potential structural alterations associated with the mutation.

## **Results**

### *Clinical Characteristics*

In total, 114 subjects were included in the current study (2023-2025, Kermanshah, Iran),

based on the reasons for referral to the Medical Genetic Lab by a medical specialist. According to table 1, the minimum, maximum, and mean ages among these cases were 6 months, 58 years, and 18.82±13.15 years, respectively. The sex distribution was nearly balanced, with 49.12% male and 50.88% female. Among 114 cases, 56.14% were born to consanguineous parents. Another important factor was the zygosity status of patients, which can help assess the severity of Autosomal Recessive (AR) status. Accordingly, 41% of 139 mutations showed homozygous genotyping, 56.84% indicated heterozygous genotyping, and only a small proportion was related to hemizygous genotyping (2.16%). Of note, we observed five inheritance patterns, including AR (69.17%), AD (20.83%), AD/AR (7.5%), XLR (0.84%), and XLD (1.66%).

### *WES Test Results*

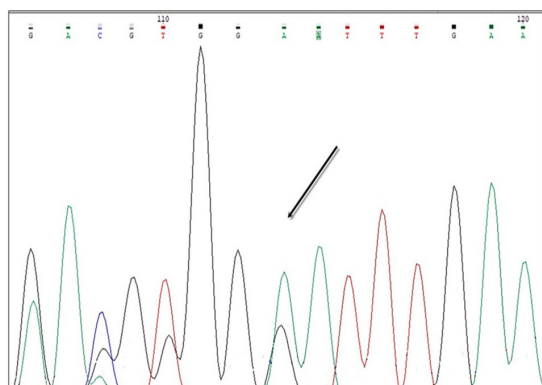
WES results were analyzed by validated NGS-based databases including Franklin and Varsome. Notably, of the 139 mutations, 73 novel mutations (52.52%) (supplementary table 1) were found in 56 cases; among them, 31 cases (55.36%) were born to consanguineous parents. Therefore, half of the novel mutations might be related to parents' relatedness status (table 1). After removing duplicate genes, we identified 69 unique genes. To classify these genes, we conducted a PPI analysis that revealed four categories. Each category contained several protein-coding genes, including Group 1 with 19 members, Group 2 with 10 members, Group 3 with five members, and Group 4 with four members. We selected the largest category (Group 1) for further analysis. Thus, the candidate gene list of the present study consists of 19 protein-coding genes as follows: *ABCC2*, *ALDH4A1*, *ASH1L*, *ATP7B*, *CFTR*, *CUL7*, *DDX3X*, *EP300*, *GATAD2B*, *GPT2*, *IVD*, *KDM4B*, *KRT18*, *MMUT*, *POLR3B*, *RALGAP1*, *SPR*, *USP7*, and *WDR26*. Notably, the related novel mutations are summarized in Table 3. We successfully carried out a primary and a deep *in silico* (two-step) strategy to identify potential interactions, possible biomarkers, and highly significant phenotypes resulting from the interplays among the candidate gene list. In the first step, we performed PPIs, GMIs, TF-gene Interactions (TFGIs), protein-drug interactions (PDIs), and PCIs using the STRING-MODEL and NetworkAnalyst tools. In the second step, we designed an EA and a CEA for the candidate gene list, leading to deeper insight into new findings and reliable predictions.

**Table 3:** Details of novel mutations reported in the current study according to the candidate group of genes involved in intellectual disabilities

| Gene    | Variant                                                                       | Classification | Zygosity | Inheritance | Disease                                                                                              | OMIM   | AF       |
|---------|-------------------------------------------------------------------------------|----------------|----------|-------------|------------------------------------------------------------------------------------------------------|--------|----------|
| ABCC2   | chr10:101591480:G>A:NM_000392:exon22:c.G2996A:p.W999X                         | LP             | Het      | AR          | Dubin-Johnson syndrome                                                                               | 237500 | NA       |
| ALDH4A1 | chr1:19212978:T>C:NM_003748.4:exon4:c.277A>G:p.K93E                           | VUS            | Het      | AR          | Hyperprolinemia, type II                                                                             | 239510 | 0        |
| ASH1L   | hg38:chr1-155478711 A>G                                                       | VUS            | Het      | AD          | Intellectual developmental disorder, autosomal dominant 52                                           | 617796 | 2.00E-06 |
| ATP7B   | chr13:52548742:C>T:ATP7B:NM_00053:exon2:c.G614A:p.G205E                       | LP             | Hom      | AR          | Wilson disease                                                                                       | 277900 | NA       |
| CFTR    | chr7:117199646:NM_000492.4:CTTT>C:p.Phe508del*                                | LP             | Het      | AR          | Cystic fibrosis                                                                                      | 219700 | NA       |
| CUL7    | Chr6:43017767:GATGA>G:ENST00000535468.1:exon6                                 | LP             | Het      | AR          | 3-M syndrome 1                                                                                       | 273750 | NA       |
| DDX3X   | chrX:41203491:G>A:NM_001356:exon10:c.865-1G>A                                 | LP             | Het      | XLD         | Intellectual developmental disorder, X-linked syndromic, Snijders-Blok type                          | 300958 | NA       |
| EP300   | Chr22:41566433:C>A:NM_001429:exon27:c.4310C>A:p.Ala1437Glu                    | LP (de novo)   | Het      | AD          | Rubinstein-Taybi syndrome 2                                                                          | 6E+05  | NA       |
| GATAD2B | chr1:153788776 C>A                                                            | VUS            | Het      | AD          | GAND syndrome                                                                                        | 615074 | NA       |
| GPT2    | chr16:46884738: TC>T                                                          | LP             | Hom      | AR          | Neurodevelopmental disorder with microcephaly and spastic paraplegia                                 | 616281 | NA       |
| IVD     | chr15:40707138:GGGCTGGACCTGGAGC>G:NM_002225:exon8:c.842_856del:p.D281_L285del | VUS            | Het      | AR          | Isovaleric acidemia                                                                                  | 243500 | NA       |
| KDM4B   | chr19-5047487 G>A                                                             | strong VUS     | Het      | AD          | Intellectual developmental disorder, autosomal dominant 65                                           | 619320 | NA       |
| KRT18   | chr12:53343131:-:A:NM_000224.3:exon1:c.174_175insA:p.Gly59ArgfsTer97          | LP             | Het      | AR          | Cirrhosis, cryptogenic                                                                               | 215600 | NA       |
| MMUT    | chr6:49426962:T>C:NM_000255:exon2:c.218A>G:p.Tyr73Cys                         | LP             | Het      | AR          | Methylmalonic aciduria, mut(0) type                                                                  | 251000 | NA       |
| POLR3B  | chr12: 106903283: A>G: NM_018082.6: exon28:c.3358A>G:p.Met1120Val             | LP             | Het      | AD          | Charcot-Marie-Tooth disease, demyelinating, type 1I                                                  | 619742 | 1.2E-05  |
| RALGAP1 | Chr14:36194324:TC>T:NM_:194301.4:exon14:c.1771del:p.E591Nfs*4                 | LP             | Het      | AR          | Neurodevelopmental disorder with hypotonia, neonatal respiratory insufficiency, and thermoregulation | 618797 | 0        |
| RALGAP1 | Chr14:36194327:G>T:NM_:194301.4:exon14:c.1769C>A:p.T590K                      | LP             | Het      | AR          | Neurodevelopmental disorder with hypotonia, neonatal respiratory insufficiency, and thermoregulation | 618797 | 0        |
| SPR     | chr2:73115446:C>T:NM_003124:exon2:c.308C>T:p.S103F                            | LP             | Hom      | AD/AR       | Dystonia, dopa-responsive, due to sepiapterin reductase deficiency                                   | 612716 | 3E-06    |
| USP7    | hg38:chr16-8936586:C>A                                                        | VUS            | Het      | AD          | Hao-Fountain syndrome                                                                                | 616863 | 1.00E-06 |
| WDR26   | hg38: chr1-224411545 G>A                                                      | VUS            | Het      | AD          | Skraban-Deardorff syndrome                                                                           | 617616 | 1.20E-05 |

Notably, the mutation Phe508del in the CFTR gene has been previously reported and is not novel; however, the disease-causing inheritance of CFTR is AR, and we found a heterozygous patient who carried this mutation, which describes the novelty we considered. VUS: Variant of uncertain significance; LP: Likely pathogenic; P: Pathogenic; Hom: Homozygote; Het: Heterozygote; Hem: Hemizygote; AR: Autosomal recessive; AD: Autosomal dominant; XLR: X-linked recessive; XLD: X-linked dominant; AF: Allele frequency

## ATP7B-Mother



## ATP7B-Affected Fetus

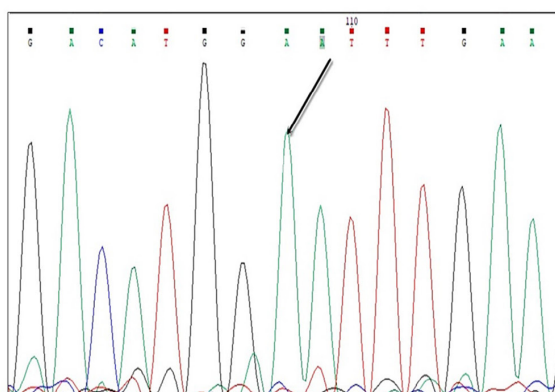


**Figure 1:** This picture represents the mutation confirmation of ATP7B (c.G614A) by Sanger sequencing in the mother and her affected fetus. The arrows show the normal (up) and mutated (down) alleles.

### Mutation Confirmation by Sanger Sequencing

As mentioned before, we checked all novel mutations in the families by Sanger sequencing, and here we present the related sequencing results for the candidate gene list of 19 members. According to table 3, 20 mutations were found in 19 genes, and we displayed the Sanger results for an important mutation (ATP7B: NM\_000053: exon2: c. G614A: p. Gly205Glu) in a mother and her affected fetus (figure 1). Remarkably, the other Sanger results were archived and will be available upon reasonable request from the corresponding author via email.

### Homology Modeling and Superimposition Results

The ATP7B protein was selected for homology modeling and superimposition analyses. Human ATP7B was found in UniProtKB with ID P35670, which was used to obtain the canonical sequence



**Figure 2:** This picture shows the superimposed model of Normal (Gly205) and Mutant (Glu205) structures of ATP7B modeled by MODELLER and compared with Matchmaker/Chimera. The normal structure is colored in cyan and the mutated structure in red. The completely superimposed parts are in green.

of the protein, comprising 1465 amino acids. A mutation was detected at residue 205<sup>th</sup>, substituting Gly (wild-type) with Glu (mutated). Among the models built by I-TASSER-MTD, SWISS-MODEL, and MODELLER, MODELLER produced the most accurate results. MODELLER PDB templates were 7XUK, 7SI3, 8Q76, and 3J09, all with identity percentages higher than 38% and query coverage higher than 65% following Blastp ([https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE\\_TYPE=BlastSearch&LINK\\_LOC=blasthome](https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome)). MODELLER built 5 models, and the best structure had the lowest Discrete Optimized Protein Energy (DOPE) score of -115491.74219, indicating the lowest entropy. This model was selected for further structural validation and superimposition. For the mutated structure, another MODELLER run built 5 models, and the best had a DOPE score of -116940.08594. The wild-type structure of ATP7B showed validation assessments, including a Ramachandran favored score of 84.07% and a z-score of -0.47.

The mutated structure passed the qualification assurance items as follows: Ramachandran of 84.69% and a z-score of -0.71. Finally, to visualize and compare the two modeled structures, Matchmaker in Chimera performed superimposition with the Needleman-Wunsch alignment algorithm and the BLOSUM-85 matrix. Figure 2 illustrates the superimposed model of wild-type and mutated ATP7B at residue 205<sup>th</sup>. There is no obvious difference between the two

modeled structures; however, other evidence strongly supports the scenario that changing Gly to Glu may alter the tertiary structure of ATP7B and thus may lead to its pathogenicity, which in turn should be considered a computational prediction, not a functional finding. One key clue is that Gly is neutral and Glu is negatively charged, and another is that validation assessments of both structures showed minor differences in DOPE, Ramachandran, and z-scores.

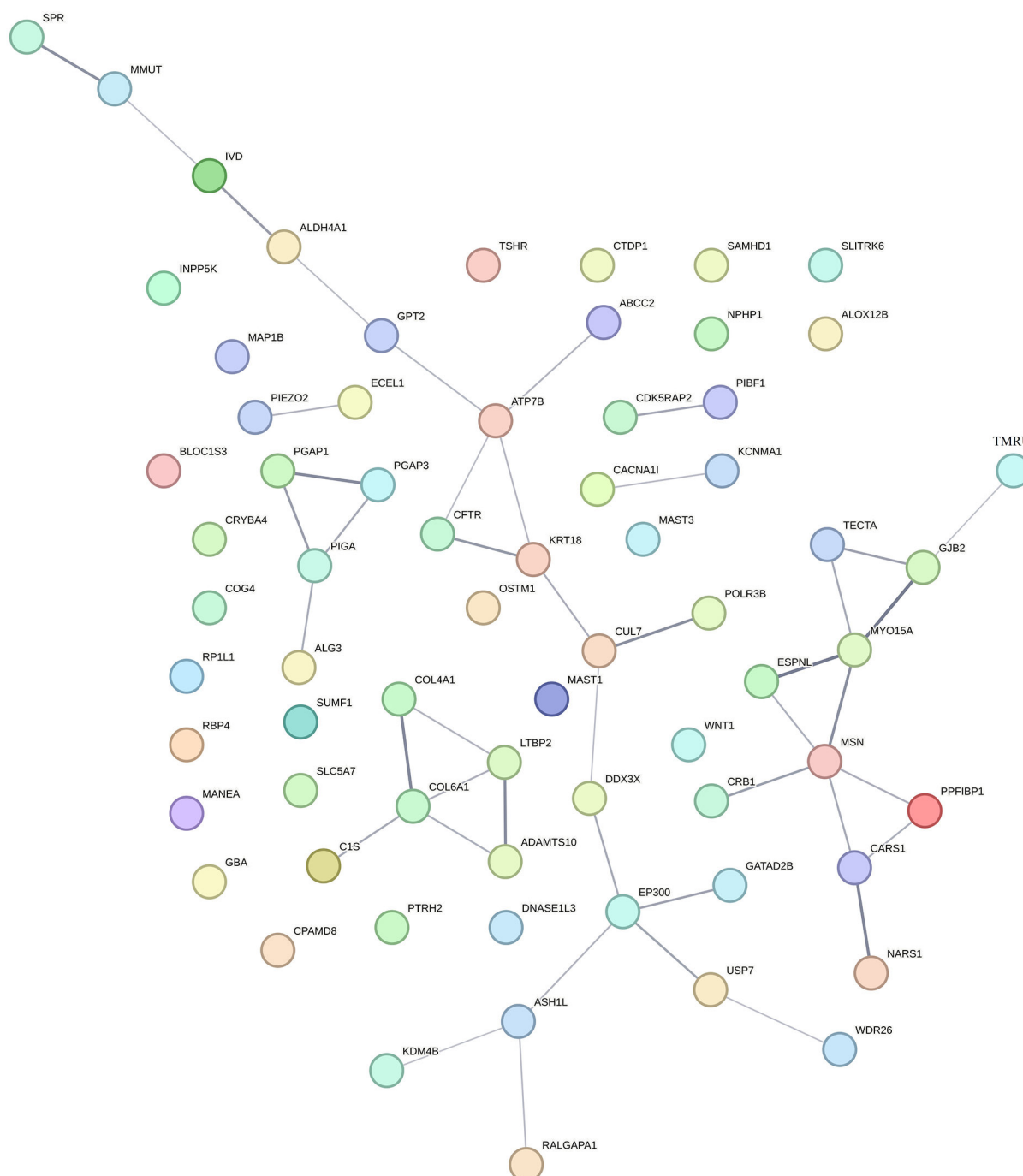


**Figure 3:** This picture highlights the STRING model of all genes found by Whole-Exome Sequencing (WES), which showed important mutations due to the lack of clinical data, representing the 4 connected Groups with various protein-coding genes as members. As is obvious, ATP7B is the most interacted member in the candidate group. The statistical details for this network were: average node degree=1.29, average node degree=0.427, and PPI enrichment P=0.00015.

### Protein-Protein Interaction (PPI) Results

As mentioned before, using the STRING-MODEL tool, we started with a primary PPI network containing all initial genes (69 genes) from clinical exome analysis (figure 3) and then selected the largest network (19 genes) (figure 4). ATP7B is the most connected member of the candidate group (Group 1), with 4 interactions with ABCC2, GPT2, CFTR, and KRT18.

### Gene-miRNA Interaction (GMIs) Results

Applying NetworkAnalyst and the miRTarBase v8.0 database, GMIs were predicted using a Bi/Tripartite Linear model for the candidate genes. GMI results showed that hsa-miR-615-3p was the most connected miRNA in this GMI, followed by hsa-miR-26a-5p and hsa-let-7b-5p, each with 4 connections (figure 5).

### TF-Gene Interactions (TFGIs) Results

Using NetworkAnalyst and adjusting the ENCODE database, TFGIs for the candidate genes were visualized in a Force Atlas model. According to figure 6, the most likely TFs in this model were as follows: TGIF2 had six interactions, and CREB3L1, KLF11, and ELF1 each had five interactions.

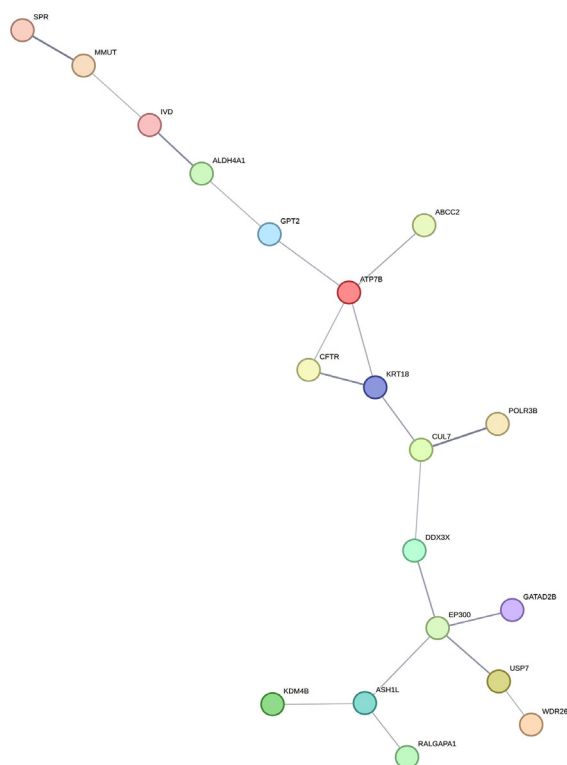


**Figure 4:** This picture indicates the Protein-Protein Interaction (PPI) network of the Group 1 candidate gene list visualized by STRING-MODEL, highlighting ATP7B, CUL7, ASH1L, and EP300 as the most connected proteins. The statistical details for this network included: average node degree=2, average node degree=0.447, and PPI enrichment P= 5.43e-05.

### Protein-Chemical Interactions (PCIs) Results

Employing NetworkAnalyst and setting the Comparative Toxicogenomics Database (CTD), PCIs were mapped to candidate genes with a Sugiyama model. In figure 7, the most connected protein was ABCC2 with 206 connections, and the most connected chemical was Valproic Acid with 16 connections, including ABCC2, CFTR, EP300, KRT18, ATP7B, KDM4B, GPT2, MUT, DDX3X, ALDH4A1, GATAD2B, POLR3B, ASH1L, CUL7, RALGAPA1, USP7, WDR26, and SPR. This indicates that the majority of the candidate list (16 of 19) is linked to valproic acid, providing additional support for our primary findings.

### Enrichment Analysis (EA) Results

To provide a clearer view of the primary clinical exome findings and the initial results from the first step of *in silico* analysis, we conducted an EA and a CEA for the candidate gene list. EA comprises three subgroups, including Pathway analysis, GO, and Diseases/Drugs Assessments (DDAs), which are summarized below and discussed in the discussion section. Notably, all categorizations are filtered by the adjusted P value (q-value) as the significance level.

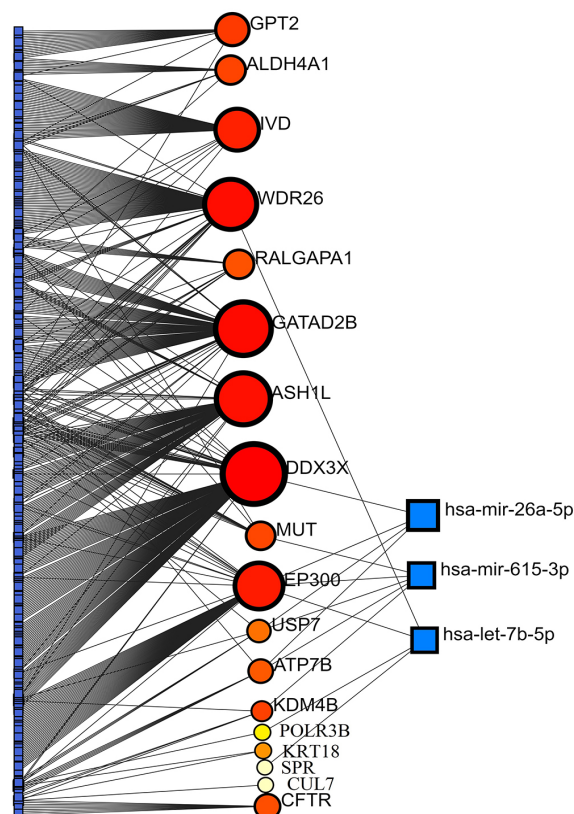


**Figure 5:** This picture represents the Gene-miRNA interactions (GMIs) network found by NetworkAnalyst, indicating another clue for the relationships among candidate genes and highlighting hsa-miR-615-3p as the most interacted miRNA.

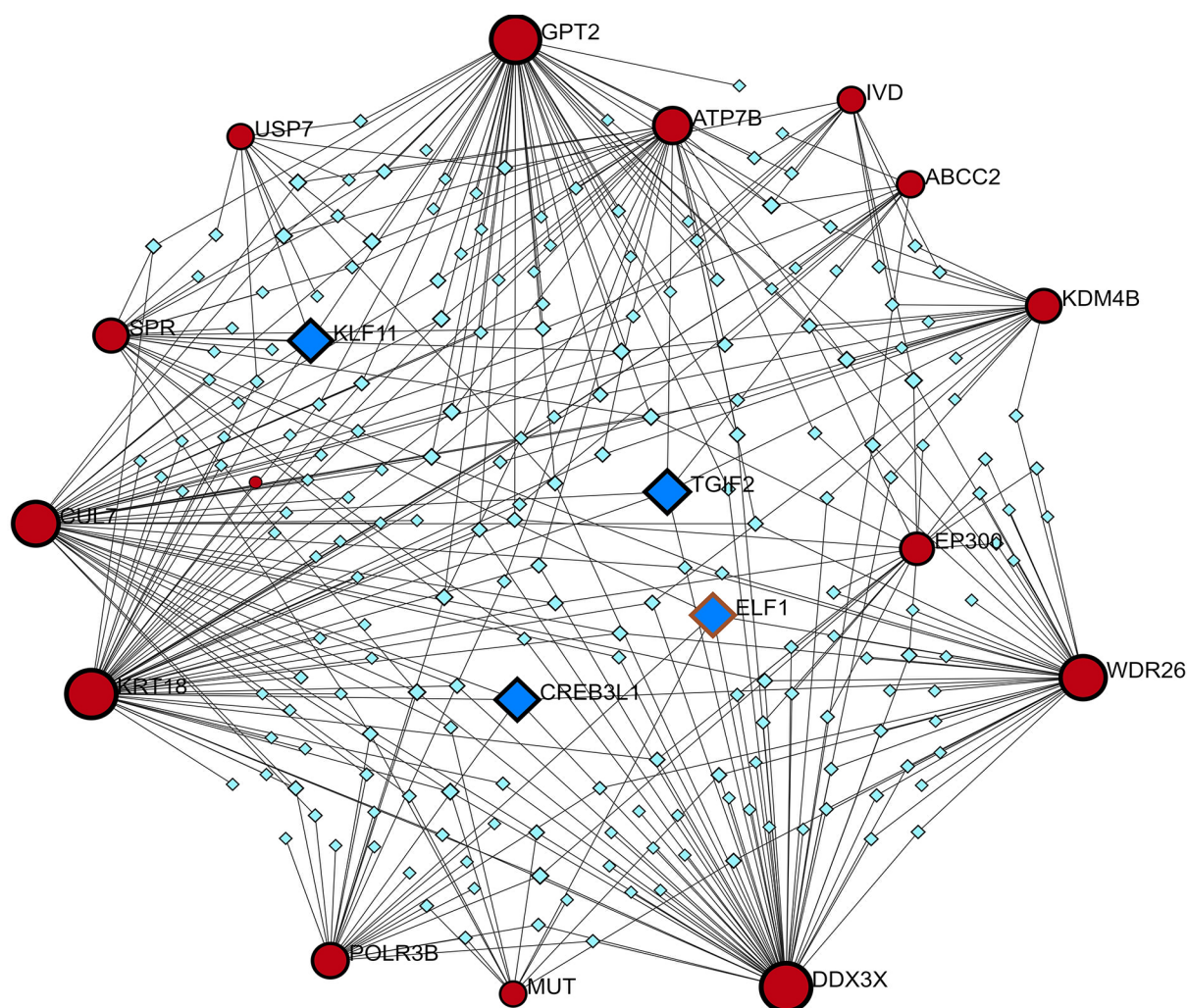


**Figure 6:** This illustration shows the Force atlas model resulting from the candidate genes, highlighting the most important Transcription Factors (TFs), including TGIF2, CREB3L1, KLF11, and ELF1.

### Pathway Analysis Results

According to the Kyoto Encyclopedia of Genes and Genomes (KEGG) and Reactome databases, pathway analysis of the candidate gene list revealed that the most significant pathway was Alanine, aspartate, and glutamate metabolism ( $P=5.6E-4$ ;  $Q=0.011$ ;  $OR=67.05$ ). Interestingly, the most significant pathway based on Reactome was Chromatin Modifying Enzymes R-HSA-3247509 ( $Q=0.012$ ;  $OR=22.5$ ), which might be a plausible scenario for Epigenetic influences among these candidate genes (table 4).

### Gene Ontology (GO) Results

GO cellular component, GO molecular function, and GO biological process are the subcategories of GO enrichment analysis. Mitochondrial matrix (GO:0005759) was the top-scoring item in the GO cellular component subcategory, with a  $P$  value of  $3.06E-4$ , a  $Q$  value of 0.012, and an  $OR$  of 14.91. Negative regulation

of gluconeogenesis (GO:0045721) was the most significant item in the GO biological process subcategory ( $P=8.912E-05$ ;  $Q=0.02326$ ; and  $OR=180.71$ ). Additionally, chromatin DNA binding (GO:0031490) was the top-scoring item in the GO molecular function subcategory, with a  $Q$  value of 0.06 and a significant  $P$  value of 0.002 (table 5).

### Diseases/Drug Assessments (DDA)

We adjusted DDA for two databases, including GeDiPNet and DisGeNET. DDA identified mental retardation (MR) ( $P=3.11E-10$ ;  $Q=8.85E-08$ ; and  $OR=22.53$ ) and ID ( $P=2.34E-08$ ;  $Q=2.5E-05$ ; and  $OR=15.22$ ) as the top-scoring significant phenotypes, reflecting the interplay among the candidate gene list (table 6).

### Clustering Enrichment Analysis (CEA) Results

Using Metascape to identify additional connections, we conducted a CEA on the candidate gene list. Figure 8 highlights the

previous EA with additional algorithms and confirms the prior predictions. Accordingly, the bar chart indicates that the most significant

process resulting from interactions among the candidate genes is the Alpha-amino acid metabolic process (GO: 1901605).

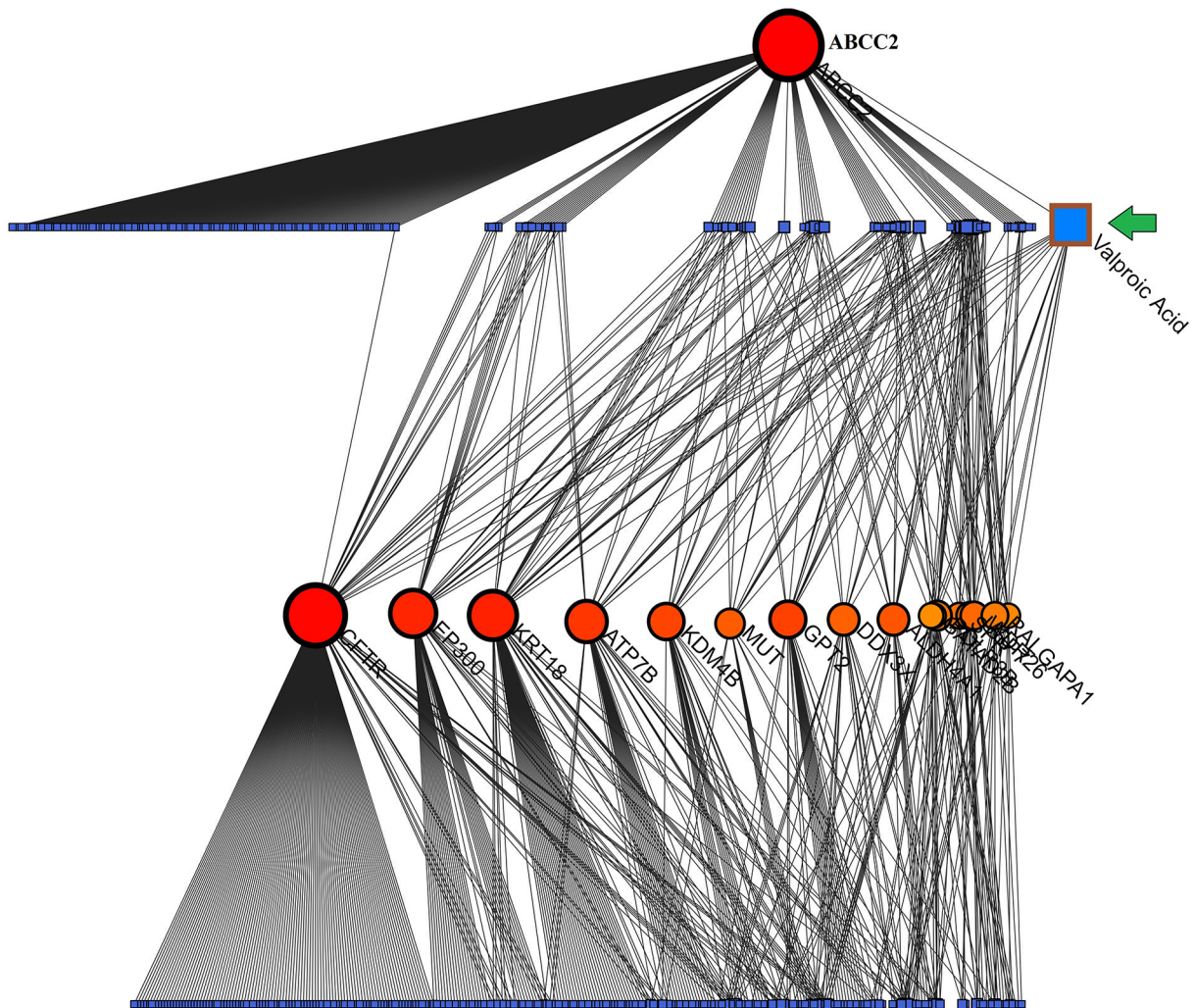


**Figure 7:** This picture visualizes the Sugiyama model of Protein-Chemical Interactions (PCIs) based on the candidate gene list. As is clear, valproic acid seems to be the most promising chemical with 16 connections.

**Table 4:** Pathway Analysis of candidate genes based on KEGG and Reactome databases

| Index    | Name                                                                  | P        | Q       | OR    |
|----------|-----------------------------------------------------------------------|----------|---------|-------|
| KEGG     | Alanine, aspartate, and glutamate metabolism                          | 0.000558 | 0.01151 | 67.05 |
| KEGG     | ABC transporters                                                      | 0.000826 | 0.01151 | 54.55 |
| KEGG     | Viral carcinogenesis                                                  | 0.000886 | 0.01151 | 18.54 |
| KEGG     | Valine, leucine, and isoleucine degradation                           | 0.00094  | 0.01151 | 50.98 |
| Reactome | Chromatin Modifying Enzymes R-HSA-3247509                             | 6.59E-05 | 0.01258 | 22.5  |
| Reactome | Regulation of TP53 Activity Through Acetylation R-HSA-6804758         | 0.000342 | 0.02669 | 86.95 |
| Reactome | Regulation of TP53 Activity R-HSA-5633007                             | 0.000419 | 0.02669 | 24.14 |
| KEGG     | Bile secretion                                                        | 0.003258 | 0.03193 | 26.59 |
| Reactome | Formation of TC-NER Pre-Incision Complex R-HSA-6781823                | 0.001102 | 0.04163 | 46.9  |
| Reactome | Cytosolic Sensors of Pathogen-Associated DNA R-HSA-1834949            | 0.001613 | 0.04163 | 38.42 |
| Reactome | Gap-filling DNA Repair Synthesis and Ligation In TC-NER R-HSA-6782210 | 0.001664 | 0.04163 | 37.8  |
| Reactome | Dual Incision in TC-NER R-HSA-6782135                                 | 0.001716 | 0.04163 | 37.2  |
| Reactome | Deubiquitination R-HSA-5688426                                        | 0.002205 | 0.04163 | 13.39 |
| Reactome | Positive Epigenetic Regulation of rRNA Expression R-HSA-5250913       | 0.002217 | 0.04163 | 32.53 |
| Reactome | ABC Transporter Disorders R-HSA-5619084                               | 0.002398 | 0.04163 | 31.23 |

P refers to P value as the significance level of lower than 0.05; Q refers to an amount of adjusted P value; and OR refers to the odds ratio.

**Table 5:** Gene ontology (GO) analysis of candidate genes based on GO biological process, GO cellular component, and GO molecular function

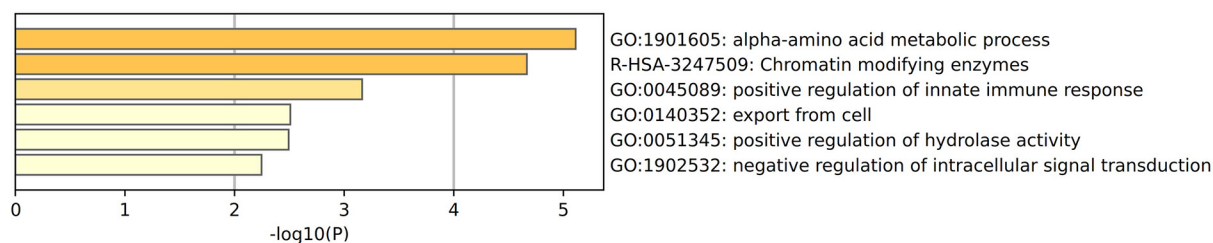
| Index                 | Name                                                                 | P         | Q       | OR     |
|-----------------------|----------------------------------------------------------------------|-----------|---------|--------|
| GO Cellular Component | Mitochondrial Matrix (GO:0005759)                                    | 0.0003063 | 0.01225 | 14.91  |
| GO Cellular Component | Intracellular Organelle Lumen (GO:0070013)                           | 0.0009995 | 0.01999 | 8.03   |
| GO Biological Process | Negative Regulation of Gluconeogenesis (GO:0045721)                  | 8.912E-05 | 0.02326 | 180.71 |
| GO Biological Process | Negative Regulation of Cellular Biosynthetic Process (GO:0031327)    | 0.0002812 | 0.02326 | 27.77  |
| GO Biological Process | Alpha-Amino Acid Catabolic Process (GO:1901606)                      | 0.0003419 | 0.02326 | 86.95  |
| GO Biological Process | Negative Regulation of Carbohydrate Metabolic Process (GO:0045912)   | 0.0003661 | 0.02326 | 83.84  |
| GO Biological Process | Regulation Of Gluconeogenesis (GO:0006111)                           | 0.0003911 | 0.02326 | 80.94  |
| GO Biological Process | Negative Regulation of Small Molecule Metabolic Process (GO:0062014) | 0.0004436 | 0.02326 | 75.71  |
| GO Biological Process | Positive Regulation of Interferon-Beta Production (GO:0032728)       | 0.0005284 | 0.02326 | 69.02  |
| GO Biological Process | Amino Acid Catabolic Process (GO:0009063)                            | 0.0005889 | 0.02326 | 65.18  |
| GO Biological Process | Regulation Of Signal Transduction by P53 Class Mediator (GO:1901796) | 0.000901  | 0.03164 | 52.12  |
| GO Biological Process | Regulation Of Interferon-Beta Production (GO:0032648)                | 0.001019  | 0.03221 | 48.86  |
| GO Molecular Function | Chromatin DNA Binding (GO:0031490)                                   | 0.002779  | 0.06131 | 28.9   |

P refers to the P value, which is considered as the amount lower than 0.05; Q means adjusted P; and OR refers to the odds ratio.

**Table 6:** Diseases/drug assessments (DDA) of candidate genes based on DisGeNET and GeDiPNet databases

| Index    | Name                              | P        | Q        | OR     |
|----------|-----------------------------------|----------|----------|--------|
| GeDiPNet | Mental Retardation                | 3.11E-10 | 8.85E-08 | 22.53  |
| DisGeNET | Intellectual Disability           | 2.34E-08 | 2.5E-05  | 15.22  |
| GeDiPNet | Developmental Delay               | 3.28E-07 | 4.68E-05 | 13.48  |
| DisGeNET | Sialorrhea                        | 6.5E-06  | 0.002321 | 103.88 |
| DisGeNET | Droling                           | 6.5E-06  | 0.002321 | 103.88 |
| DisGeNET | Hepatitis                         | 9.39E-06 | 0.002514 | 22.51  |
| DisGeNET | Liver diseases                    | 1.46E-05 | 0.002784 | 14.91  |
| DisGeNET | Range of joint movement increased | 1.56E-05 | 0.002784 | 76.27  |
| GeDiPNet | Carcinoma Of the Head and Neck    | 4.16E-05 | 0.003148 | 54.11  |
| GeDiPNet | Movement Disorders                | 4.94E-05 | 0.003148 | 24.29  |
| GeDiPNet | Dysarthria                        | 6.21E-05 | 0.003148 | 14.99  |
| GeDiPNet | Hyperglycinuria                   | 6.63E-05 | 0.003148 | 213.58 |
| DisGeNET | Mental deficiency                 | 2.78E-05 | 0.004255 | 10.79  |
| GeDiPNet | Sleep Apnea                       | 0.00023  | 0.007419 | 29.78  |
| GeDiPNet | Dysmorphic Features               | 0.000251 | 0.007419 | 15.73  |
| GeDiPNet | Esophageal Varix                  | 0.000253 | 0.007419 | 102.09 |
| GeDiPNet | Autism                            | 0.000284 | 0.007419 | 10.71  |
| DisGeNET | Glycinuria                        | 7.73E-05 | 0.008386 | 195.77 |
| DisGeNET | Cholestasis in newborn            | 8.65E-05 | 0.008386 | 41.91  |
| DisGeNET | Genetic Diseases, Inborn          | 8.89E-05 | 0.008386 | 13.86  |

P, Q, and OR stand for P value (significance level lower than 0.05), adjusted P, and Odds ratio.

**Figure 8:** This picture represents the bar chart for candidate genes based on Clustering enrichment analysis (CEA) employing the Metascape tool.

## Discussion

Results from both primary and secondary bioinformatics analyses highlighted the main idea and revealed valuable findings as follows:

*ATP7B* was the most promising member of the candidate group in PPIs; *hsa-miR-615-3p* was the most interacting miRNA in GMIs; *TGIF2* showed the top-connected TF in TFGIs; and valproic acid was found as the best-scored

chemical in PCIs. Additionally, a systems biology assessment was successfully carried out on the candidate gene list, which showed that metabolic processes, epigenetic influences, mitochondrial matrix, and negative regulation of gluconeogenesis all led to mental retardation and ID manifestations.

As more patients are identified using exome and genome sequencing in healthcare settings or investigations, it becomes increasingly important to enable the detection of potentially curable cases at the variant or gene level.<sup>7</sup> We searched for the MeSH Terms “Metabolic” + “Intellectual disability” in the title of papers published in PubMed. Results revealed that there are 883 publications from 1998 to 2024; among them, 28 publications completely used the aforementioned terms in their titles. Notably, 22 were original papers (research articles, commentaries, editorials, multicenter studies, comparative studies, clinical trials, and observational studies), four were reviews and meta-analyses, and two were case reports.

Notwithstanding their potential, whole exome and whole genome sequencing present multiple limitations and obstacles when used in therapeutic settings. The key problems concerning genome-wide sequencing in clinical settings are data analysis, results reporting, and conclusion sharing.<sup>15</sup> Genetic understanding is advancing by the day, yet it remains elementary. While a distinct variation in the DNA sequence may be discovered, the relationship underlying that variation and the disease is often unclear. Whole genome sequencing produces far larger collections of VUS than whole exome sequencing.<sup>15</sup> Clinicians have highlighted the single most significant problem as interpreting and responding to VUS.<sup>15, 16</sup> In accordance with the American College of Medical Genetics and Genomics (ACMG) criteria.<sup>17</sup> A variant can be categorized as pathogenic (disease-causing), likely pathogenic, VUS, likely benign, or benign based on what is currently known about that variation (i.e., reported in the literature) or what has been identified in family members. Furthermore, bioinformatics tools employed for managing data, filtering, and classifying variants are based on incomplete databases, which means they contain not fully finalized algorithms that are always being refined as new knowledge is gained.<sup>15</sup> Analysis of an individual genome necessitates substantial processing capacity and electronic space for storage, which neither individual nor institution can always provide; hence, connections of national and international groups are growing for sharing efficient procedures and experience.<sup>18</sup>

It has recently been demonstrated in some countries that ES, when used as the first-line test for the detection of uncommon genetic diseases, may be performed at a fraction of the cost and triple the diagnostic rate in children suffering from suspected monogenic disorders.<sup>19, 20</sup> It is challenging to apply the initially reported results to different contexts; nevertheless, due to the heterogeneity in clinical manifestations, the number of samples, and the health economic assessment criteria.<sup>21</sup> The economic effect of obtaining a molecular detection in ID extends beyond the patient and their family to the healthcare system, as genetic counseling plays a secondary role in preventing future cases. Similar to this, a diagnosis is beneficial for social and educational contexts.<sup>22, 23</sup> In addition to helping diagnose or estimate the risk of disorders that might otherwise go undetected using clinical history, physical checkup, and metabolic or biochemical testing, GWAS, such as WES or WGS, may identify whether an error is included.<sup>18</sup>

Gene panels leverage NGS technology to decode the genome and to evaluate and interpret the set of selected “core genes”. The panels’ candidate genes are selected to make it easier to identify a certain group of illnesses, syndromes, or phenotypes. They typically contain the genes known to cause the disease, though VUS are occasionally chosen as well. Gene panels come in two varieties: virtual panels and targeted gene panels (TGP). TGP, also known as “partial exome testing,” concentrates on examining the candidate genes while omitting to examine the entire genome.<sup>24</sup> Conversely, virtual gene panels analyze the relevant genes *in silico* by using ES or GS libraries. Thus, virtual gene panels enable further repetitive analysis.<sup>25</sup> Selecting the appropriate genes for the panels requires both clinical experience in DD/ID and the pertinent genetic information. According to current research, the target list of gene panels has to be updated on a regular basis to account for newly discovered gene associations related to DD/ID.<sup>26</sup> da Cunha Leite stated that they were able to determine genotype-phenotype correlations for many genes linked to ID/GDD/MCA by applying the ID gene panel to exome data, which had excellent accuracy and specificity.<sup>27</sup>

Target exome sequencing has been shown to provide a diagnostic rate of between 21% and 55.7% in several cohorts of people with ID. Using four distinct panels, Pেকেles and colleagues studied a cohort of 48 patients and achieved a diagnosis rate of 21%.<sup>28</sup> Using a sample of 133 patients, Yamamoto and others found a 29.3% diagnostic rate.<sup>29</sup> Geldon and

her colleagues showed a diagnosis rate of 34% in a trial involving 108 patients and 4,813 genes.<sup>30</sup> Employing a panel containing 4,813 genes, Stojanovic and others found that 88 children with moderate to advanced ID / GDD had a 55.7% diagnosis rate.<sup>31</sup> Another study based on WES or WGS suggested a diagnostic rate ranging between 8% and 60%, depending on the selection criteria.<sup>32</sup> From a diagnostic perspective, Martínez and colleagues claimed that the target gene panel and WES/WGS have comparable diagnostic yields, making them comparable approaches.<sup>33</sup>

There are controversial reports about being the first-tier test between Targeted Gene Panels and Exome gene panels.<sup>24, 34</sup> However, due to several main reasons discussed in the following, exome gene panels are rationally more important: 1) The most well-recognized guideline in the field of NGS, ACMG, has released an updated version that suggests ES and GS as the first-tier genetic screening for patients with ID.<sup>35</sup> 2) One major disadvantage of focused NGS panels is that they might be outdated swiftly. While a panel is being designed and verified for clinical use, new research is being published that discovers newly described disease genes. In fact, 23% of positive WES findings come from genes defined within the previous two years, whereas 7% are novel gene findings.<sup>36</sup> Since any alterations in both well-characterized and novel genes may be analyzed and interpreted, one benefit of WES is the capacity to sequence the complete exome at once. This also allows for re-interpretation in the event that new genetic correlations are discovered. Aside from these benefits, an ES approach can also be used to sequence the exomes of several family members at once in labs offering sequencing of trios, detecting novel genes, and analyzing a much higher number of genes at reasonable prices.<sup>37</sup>

There are some limitations in the current study, such as functional validation of new mutations, as well as the sample size and using WES rather than Whole-genome sequencing (WGS) due to the higher prices. Additionally, functional validation (*in vitro/in vivo*) and molecular dynamics simulation for ATP7B were not done. Finally, geographic bias (single center, western Iran) and lack of long-term clinical outcomes were other limitations.

## Conclusion

The current study extracted suspicious cases from all WES results of patients referred to our medical genetics lab over two years and identified 114 cases. Of these, 73 harbored

VUS, likely pathogenic, or pathogenic mutations with no definitive clinical reports yet. This report presented a notable number of new mutations and employed a distinctive *in silico* analysis strategy. Accordingly, this research proposed that metabolic processes may influence the impact of new mutations, resulting in ID/DD manifestations, and provided evidence-based bioinformatics clues. Moreover, this report highlighted the importance of NGS gene panels for clinical exome sequencing as the first-tier detection in cases with unknown causes of ID/DD signs, rather than biochemical/metabolic tests and targeted gene panels.

## Acknowledgment

We want to express our sincere appreciation to all the patients and their families for their participation and to all of our colleagues in Dr. Shaveisi-zadeh Medical Genetic Lab and Islamic Azad University North Tehran Branch. This paper is derived from the Ph.D. thesis of Vahid Omarmeli, which has no grant.

## Authors' Contribution

V.O: Data acquisition, analysis, interpretation, and drafting; E.S: Conceptualization and reviewing the manuscript; M.P: Data interpretation and reviewing the manuscript; S.M: Data interpretation and reviewing the manuscript; M.M: Data interpretation and reviewing the manuscript; All authors gave their final approval of the version to be published and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

## Declaration of AI

The authors declare that no AI tools were used in the preparation of this manuscript.

**Conflict of Interest:** None declared.

## References

- 1 Belanger SA, Caron J. Evaluation of the child with global developmental delay and intellectual disability. *Paediatr Child Health*. 2018;23:403-19. doi: 10.1093/pch/pxy093. PubMed PMID: 30919832; PubMed Central PMCID: PMC6234423.
- 2 Mithyantha R, Kneen R, McCann E, Gladstone M. Current evidence-based recommendations on investigating children

- with global developmental delay. *Arch Dis Child*. 2017;102:1071-6. doi: 10.1136/archdischild-2016-311271. PubMed PMID: 29054862; PubMed Central PMCID: PMC5738593.
- 3 Lee JS, Hwang H, Kim SY, Kim KJ, Choi JS, Woo MJ, et al. Chromosomal Microarray With Clinical Diagnostic Utility in Children With Developmental Delay or Intellectual Disability. *Ann Lab Med*. 2018;38:473-80. doi: 10.3343/alm.2018.38.5.473. PubMed PMID: 29797819; PubMed Central PMCID: PMC5973923.
  - 4 Warmerdam HAG, Termeulen-Ferreira EA, Tseng LA, Lee JY, van Eeghen AM, Ferreira CR, et al. A Scoping Review of Inborn Errors of Metabolism Causing Progressive Intellectual and Neurologic Deterioration (PIND). *Front Neurol*. 2019;10:1369. doi: 10.3389/fneur.2019.01369. PubMed PMID: 32132962; PubMed Central PMCID: PMC7040240.
  - 5 Carter MT, Srour M, Au PB, Buhas D, Dyack S, Eaton A, et al. Genetic and metabolic investigations for neurodevelopmental disorders: position statement of the Canadian College of Medical Geneticists (CCMG). *J Med Genet*. 2023;60:523-32. doi: 10.1136/jmg-2022-108962. PubMed PMID: 36822643; PubMed Central PMCID: PMC10313961.
  - 6 Shi Y, Li X, Ju D, Li Y, Zhang X, Zhang Y. Abnormal chromosomes identification using chromosomal microarray. *J Obstet Gynaecol*. 2022;42:2025-32. doi: 10.1080/01443615.2022.2074786. PubMed PMID: 35659171.
  - 7 Hoytema van Konijnenburg EMM, Wortmann SB, Koelewijn MJ, Tseng LA, Houben R, Stockler-Ipsiroglu S, et al. Treatable inherited metabolic disorders causing intellectual disability: 2021 review and digital app. *Orphanet J Rare Dis*. 2021;16:170. doi: 10.1186/s13023-021-01727-2. PubMed PMID: 33845862; PubMed Central PMCID: PMC8042729.
  - 8 Chen B, Zhan Y, Kessi M, Chen S, Xiong J, Deng X, et al. Urine Organic Acids as Metabolic Indicators for Global Developmental Delay/Intellectual Disability in Chinese Children. *Front Mol Biosci*. 2021;8:792319. doi: 10.3389/fmolb.2021.792319. PubMed PMID: 35036412; PubMed Central PMCID: PMC8757376.
  - 9 van Karnebeek CD, Stockler S. Treatable inborn errors of metabolism causing intellectual disability: a systematic literature review. *Mol Genet Metab*. 2012;105:368-81. doi: 10.1016/j.ymgme.2011.11.191. PubMed PMID: 22212131.
  - 10 Smith HS, Swint JM, Lalani SR, Yamal JM, de Oliveira Otto MC, Castellanos S, et al. Clinical Application of Genome and Exome Sequencing as a Diagnostic Tool for Pediatric Patients: a Scoping Review of the Literature. *Genet Med*. 2019;21:3-16. doi: 10.1038/s41436-018-0024-6. PubMed PMID: 29760485.
  - 11 Vallance H, Sinclair G, Rakic B, Stockler-Ipsiroglu S. Diagnostic yield from routine metabolic screening tests in evaluation of global developmental delay and intellectual disability. *Paediatr Child Health*. 2021;26:344-8. doi: 10.1093/pch/pxaa112. PubMed PMID: 34676012; PubMed Central PMCID: PMC8522830.
  - 12 Zar JH. *Biostatistical analysis*. Bangalore: Pearson Education India; 1999.
  - 13 Hochberg Y, Benjamini Y. More powerful procedures for multiple significance testing. *Stat Med*. 1990;9:811-8. doi: 10.1002/sim.4780090710. PubMed PMID: 2218183.
  - 14 Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960;20:2746. doi: 10.1177/001316446002000104.
  - 15 Bertier G, Hetu M, Joly Y. Unsolved challenges of clinical whole-exome sequencing: a systematic literature review of end-users' views. *BMC Med Genomics*. 2016;9:52. doi: 10.1186/s12920-016-0213-6. PubMed PMID: 27514372; PubMed Central PMCID: PMC4982236.
  - 16 Kohane IS, Shendure J. What's a Genome Worth? *Sci Transl Med*. 2012;4:133fs13. doi: 10.1126/scitranslmed.3004208. PubMed PMID: 22572879.
  - 17 Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17:405-24. doi: 10.1038/gim.2015.30. PubMed PMID: 25741868; PubMed Central PMCID: PMC4544753.
  - 18 Ontario H. Genome-Wide Sequencing for Unexplained Developmental Disabilities or Multiple Congenital Anomalies: A Health Technology Assessment. *Ont Health Technol Assess Ser*. 2020;20:1-178. PubMed PMID: 32194879; PubMed Central PMCID: PMC7080457.
  - 19 Stark Z, Schofield D, Alam K, Wilson W, Mupfeki N, Macciocca I, et al. Prospective comparison of the cost-effectiveness of clinical whole-exome sequencing with that of usual

- care overwhelmingly supports early use and reimbursement. *Genet Med.* 2017;19:867-74. doi: 10.1038/gim.2016.221. PubMed PMID: 28125081.
- 20 Schofield D, Rynehart L, Shrestha R, White SM, Stark Z. Long-term economic impacts of exome sequencing for suspected monogenic disorders: diagnosis, management, and reproductive outcomes. *Genet Med.* 2019;21:2586-93. doi: 10.1038/s41436-019-0534-x. PubMed PMID: 31110331.
- 21 Schwarze K, Buchanan J, Taylor JC, Wordsworth S. Are whole-exome and whole-genome sequencing approaches cost-effective? A systematic review of the literature. *Genet Med.* 2018;20:1122-30. doi: 10.1038/gim.2017.247. PubMed PMID: 29446766.
- 22 Moeschler JB, Shevell M, Committee on G. Comprehensive evaluation of the child with intellectual disability or global developmental delays. *Pediatrics.* 2014;134:e903-18. doi: 10.1542/peds.2014-1839. PubMed PMID: 25157020; PubMed Central PMCID: PMC9923626.
- 23 van Karnebeek CDM. Evaluation of the Child With Developmental Impairments. *Continuum (Minneapolis Minn).* 2018;24:228-47. doi: 10.1212/CON.0000000000000564. PubMed PMID: 29432245.
- 24 Lee IH, Lin Y, Alvarez WJ, Hernandez-Ferrer C, Mandl KD, Kong SW. WEScover: selection between clinical whole exome sequencing and gene panel testing. *BMC Bioinformatics.* 2021;22:259. doi: 10.1186/s12859-021-04178-5. PubMed PMID: 34016036; PubMed Central PMCID: PMC8139020.
- 25 Ko MH, Chen HJ. Genome-Wide Sequencing Modalities for Children with Unexplained Global Developmental Delay and Intellectual Disabilities-A Narrative Review. *Children (Basel).* 2023;10. doi: 10.3390/children10030501. PubMed PMID: 36980059; PubMed Central PMCID: PMC10047410.
- 26 Winkler EC, Wiemann S. Findings made in gene panel to whole genome sequencing: data, knowledge, ethics - and consequences? *Expert Rev Mol Diagn.* 2016;16:1259-70. doi: 10.1080/14737159.2016.1212662. PubMed PMID: 27409399.
- 27 Leite A, Pinto IP, Leijsten N, Ruitkamp-Versteeg M, Pfundt R, de Leeuw N, et al. Diagnostic yield of patients with undiagnosed intellectual disability, global developmental delay and multiples congenital anomalies using karyotype, microarray analysis, whole exome sequencing from Central Brazil. *PLoS One.* 2022;17:e0266493. doi: 10.1371/journal.pone.0266493. PubMed PMID: 35390071; PubMed Central PMCID: PMC8989190.
- 28 Pেকেles H, Accogli A, Boudrahem-Addour N, Russell L, Parente F, Srouf M. Diagnostic Yield of Intellectual Disability Gene Panels. *Pediatr Neurol.* 2019;92:32-6. doi: 10.1016/j.pediatrneurol.2018.11.005. PubMed PMID: 30581057.
- 29 Yamamoto T, Imaizumi T, Yamamoto-Shimajima K, Lu Y, Yanagishita T, Shimada S, et al. Genomic backgrounds of Japanese patients with undiagnosed neurodevelopmental disorders. *Brain Dev.* 2019;41:776-82. doi: 10.1016/j.braindev.2019.05.007. PubMed PMID: 31171384.
- 30 Gieldon L, Mackenroth L, Kahlert AK, Lemke JR, Pörmann J, Schallner J, et al. Diagnostic value of partial exome sequencing in developmental disorders. *PLoS One.* 2018;13:e0201041. doi: 10.1371/journal.pone.0201041. PubMed PMID: 30091983; PubMed Central PMCID: PMC6084857.
- 31 Stojanovic JR, Miletic A, Peterlin B, Maver A, Mijovic M, Borlja N, et al. Diagnostic and Clinical Utility of Clinical Exome Sequencing in Children With Moderate and Severe Global Developmental Delay / Intellectual Disability. *J Child Neurol.* 2020;35:116-31. doi: 10.1177/0883073819879835. PubMed PMID: 31623504.
- 32 Lindstrand A, Eisfeldt J, Pettersson M, Carvalho CMB, Kvarnung M, Grigelioniene G, et al. From cytogenetics to cytogenomics: whole-genome sequencing as a first-line test comprehensively captures the diverse spectrum of disease-causing genetic variation underlying intellectual disability. *Genome Med.* 2019;11:68. doi: 10.1186/s13073-019-0675-1. PubMed PMID: 31694722; PubMed Central PMCID: PMC6836550.
- 33 Martinez F, Caro-Llopis A, Rosello M, Oltra S, Mayo S, Monfort S, et al. High diagnostic yield of syndromic intellectual disability by targeted next-generation sequencing. *J Med Genet.* 2017;54:87-92. doi: 10.1136/jmedgenet-2016-103964. PubMed PMID: 27620904.
- 34 Soden SE, Saunders CJ, Willig LK, Farrow EG, Smith LD, Petrikin JE, et al. Effectiveness of exome and genome sequencing guided by acuity of illness for diagnosis of neurodevelopmental disorders. *Sci Transl Med.* 2014;6:265ra168. doi: 10.1126/scitranslmed.3010076. PubMed PMID: 25473036; PubMed Central PMCID: PMC4286868.
- 35 Manickam K, McClain MR, Demmer LA, Biswas S, Kearney HM, Malinowski J, et al. Exome and genome sequencing for pediatric

- patients with congenital anomalies or intellectual disability: an evidence-based clinical guideline of the American College of Medical Genetics and Genomics (ACMG). *Genet Med.* 2021;23:2029-37. doi: 10.1038/s41436-021-01242-6. PubMed PMID: 34211152.
- 36 Farwell KD, Shahmirzadi L, El-Khechen D, Powis Z, Chao EC, Tippin Davis B, et al. Enhanced utility of family-centered diagnostic exome sequencing with inheritance model-based analysis: results from 500 unselected families with undiagnosed genetic conditions. *Genet Med.* 2015;17:578-86. doi: 10.1038/gim.2014.154. PubMed PMID: 25356970.
- 37 LaDuca H, Farwell KD, Vuong H, Lu HM, Mu W, Shahmirzadi L, et al. Exome sequencing covers >98% of mutations identified on targeted next generation sequencing panels. *PLoS One.* 2017;12:e0170843. doi: 10.1371/journal.pone.0170843. PubMed PMID: 28152038; PubMed Central PMCID: PMC5289469