Genetic characterization of GSD I in Serbian population revealed unexpectedly high incidence of GSD Ib and three novel *SLC37A4* variants

Anita Skakic[1], Maja Djordjevic[2,3], Adrijan Sarajlija[2,3], Kristel Klaassen[1], Natasa Tosic[1], Bozica Kecman[2], Milena Ugrin[1], Vesna Spasovski[1], Sonja Pavlovic[1], Maja Stojiljkovic[1]

[1]*Institute of Molecular Genetics and Genetic Engineering, University of Belgrade*
[2]*Mother and Child Health Care Institute of Serbia „Dr Vukan Cupic“*
[3]*School of Medicine, University of Belgrade*

Corresponding author: Maja Stojiljkovic, PhD
Institute of Molecular Genetics and Genetic Engineering, University of Belgrade
Vojvode Stepe 444a, 11010 Belgrade, SERBIA
Tel: +381-64-2202-373, Fax: +381-11-3975-808
E-mail: maja.stojiljkovic@imgge.bg.ac.rs

Ethical approval

This study has been approved by the Ethics Committee of the Mother and Child Health Care Institute of Serbia „Dr Vukan Cupic“ in Belgrade, Serbia and has therefore been performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments. Informed consent was obtained from all individual participants included in the study.

Conflict of Interest

The authors declare that they have no conflict of interest.

Acknowledgements

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/cge.13093

This work has been funded by grants from the MESTD, Republic of Serbia (III41004) and European Commission (EU-FP7-REGPOT-316088).

ABSTRACT

Glycogen storage disease (GSD) type I is inborn metabolic disease characterized by accumulation of glycogen in multiple organs. We analyzed 38 patients with clinical suspicion of GSD I using Sanger and next-generation sequencing (NGS). We identified 28 GSD Ib and five Ia patients. In five patients, GSD III, VI, IX, cholesteryl-ester storage disease and Shwachman-Diamond syndrome diagnoses were set using NGS. Incidences for GSD Ia and GSD Ib were estimated at 1:172746 and 1:60461 live-births respectively. Two variants were identified in *G6PC* gene: c.247C>T (p.Arg83Cys) and c.518T>C (p.Leu173Pro). In *SLC37A4* gene, six variants were detected. Three previously reported variants c.81T>A (p.Asn27Lys), c.162C>A (p.Ser54Arg) and c.1042_1043delCT (p.Leu348Valfs*53) accounted for 87% of all analyzed alleles. Computational, transcription studies and/or clinical presentation in patients confirmed pathogenic effect of three novel variants: c.248G>A (p.Gly83Glu), c.404G>A (p.Gly135Asp) and c.785G>A (p.Ser263Glyfs*33 or p.Gly262Asp). In the cohort, hepatomegaly, hypoglycaemia and failure to thrive were the most frequent presenting signs of GSD Ia, while hepatomegaly and recurrent bacterial infections were clinical hallmarks of GSD Ib. All GSD Ib patients developed neutropenia while 20.6% developed inflammatory bowel disease. Our study revealed the highest worldwide incidence of GSD Ib. Furthermore, description of three novel variants will facilitate medical genetic practice.

Keywords: clinical-exome sequencing, G6PC, glycogen storage disease, incidence, mutations, SLC37A4 (G6PT)

INTRODUCTION

Glycogen storage disease (GSD) types Ia and Ib (OMIM #232200, #232220) are autosomal recessive diseases characterized by metabolic disturbance of both gluconeogenesis and glycogenolysis (1). Variants in *G6PC* and *SLC37A4* genes cause deficiency of an enzyme glucose-6-phosphatase (G6PC) and endoplasmic reticulum transporter glucose-6-phosphate translocase (SLC37A4) leading to GSD Ia and Ib respectively. G6PC/SLC37A4 is the key complex for regulation of blood glucose levels and its deficiency results in excessive accumulation of glycogen in the liver, kidney, and intestinal mucosa (1).

Besides GSD I, there are six additional GSD types (III, 0, XI, IX, VI, and IV) that affect liver and present with hypoglycemia and hepatomegaly (2). Among them, types I, III and IX are the most prevalent (3). Distribution of GSD I is pan-ethnic with incidence reported in range from 1:33000 to 1:400000 (1, 3, 4). The majority of GSD I patients seem to be affected by the type Ia (60-80%), while GSD type Ib is the less frequent type, affecting up to 20% (1).

Herein, we reported epidemiological, molecular and clinical characteristics of one of the largest cohort of patients with GSD I from a single population.

SUBJECTS AND METHODS

We conducted genetic analysis on 38 subjects (from 37 unrelated non-consanguineous families) who presented with hepatomegaly and hypoglycaemia during childhood. Hepatomegaly was defined as the liver palpable at more than 3.5 cm below right costal margin for infants and more than 2 cm for older children. Clinical suspicion of hepatomegaly was confirmed by ultrasound exam in every subject. Hypoglycaemia was defined as plasma glucose level below 2.7 mmol/L, regardless of symptom presence. Adult patients with onset

of the disease in childhood were also included in genetic study. Patients with established diagnosis of infectious hepatitis or other non-genetic etiology were excluded from the study.

All patients were evaluated at the Mother and Child Health Care Institute of Serbia "Dr Vukan Cupic" in Belgrade, which serves as the national referral center for inherited metabolic diseases, hosts newborn screening program and harbours the only metabolic laboratory in the country. To the best of our knowledge, collected cohort represents the vast majority population of GSD I patients from Serbian population.

Genetic and computational analysis

Genomic DNA was isolated from blood using QIAamp DNA-Blood-Mini-Kit (QIAGEN, Hilden, Germany). For all patients, all coding exons of *G6PC* and *SLC37A4* genes were individually amplified and DNA sequencing was performed as previously described (5). Five patients were examined by NGS using the Clinical-Exome Sequencing TruSight One Gene Panel (Illumina, San Diego, CA, USA). PCR and NGS details are given in Supplement. Segregation analysis was performed in all families to confirm that detected variants were on different chromosomes.

The effect of novel genetic variants was analyzed using different softwares: BDGP, UMD-predictor, PolyPhen-2, SIFT, MutPred, ClustalW2. We used I-TASSER server for protein structure prediction and Swiss-Pdb Viewer (http://www.expasy.org/spdbv).

RT-PCR profile and qRT-PCR analysis

Given that liver and leukocytes express the same major *SLC37A4* transcript (6), we opted to use RNA from noninvasively obtained blood mononuclear cells for the RT-PCR and qRT-PCR analyses. The qRT-PCR was performed using Kapa Syber Fast Universal Master Mix (KAPA Biosystems, USA) and 7500 Real-Time PCR System (Applied Biosystems, Foster City, CA, USA). RT-PCR and qRT-PCR primers were designed to selectively amplify

desired *SLC37A4* transcripts. Descriptions of RT-PCR and qRT-PCR reactions are given in Supplement.

RESULTS

Genetic analysis

Sanger sequencing of 38 patients suspected to have GSD I revealed that 28 patients (from 27 unrelated/non-consanguineous families) have a defect in the *SLC37A4* gene and 5 in *G6PC* gene (86.8% detection rate). Spectrum and frequency of variants, including three novel *SLC37A4* variants (p.Gly83Glu, p.Gly135Asp and p.Ser263Glyfs*33) are given in Table 1. Twelve (44%) homozygous and 15 compound heterozygous (56%) GSD Ib patients were found, while 90% of GSD Ia patients were homozygous. In addition, five patients analysed by NGS had variants in non-GSD I genes: *AGL*, *PYGL* and *PHKA2* genes associated with GSD III, VI and IX, as well as in non-GSD associated genes, *LIPA* and *SBDS*, responsible for cholesteryl-ester storage disease and Shwachman-Diamond syndrome respectively (Table S1).

Computational predictions suggested that novel c.785G>A variant, located at the first nucleotide of the fifth *SLC37A4* exon could be splicing (p.Ser263Glyfs*33) or missense (p.Gly262Asp) variant. The RT-PCR analysis revealed a smaller band corresponding to transcript lacking exon 5 (Figure 1). qRT-PCR study revealed that r.785_870del86 *SLC37A4* transcript was expressed 60-80% more in patients with c.785G>A variant in comparison to healthy controls (Figure 1), while expression level of a full length transcript was not different.

Prediction algorithms assessed novel variants p.Gly83Glu, p.Gly135Asp and p.Gly262Asp (if analysed as a missense variant) as damaging, amino acid sequence alignment indicated absolute evolutionary conservation of affected residues among all

analysed SLC37A4 orthologs (zebrafish, chicken, mouse, gorilla and human) (Table 1), and their locations in the protein were indicated in Figure 1.

Epidemiological data

According to the number of live births and the number of diagnosed patients in the time period from 1999 to 2015, the incidence of GSD I in Serbian population was estimated at 1:44786 live births. Interestingly, GSD Ib incidence is 1:60461, with yearly incidence of this entity ranging between 1:79025 to as high as 1:21886 (Table S2). Incidence of GSD Ia was calculated at 1:172746.

It is interesting to note that GSD Ia accounted for only 26%, while GSD Ib accounted for 74% of all GSD I patients in the Serbian population in the 1999-2015 period.

Clinical data

Median age at the time of diagnosis was 6 months for the patients with GSD Ia and 10 months for GSD Ib. Average delay between onset of symptoms and establishing diagnosis was 5,8 months for the whole group (range 0-42 months).

Hepatomegaly, hypoglycaemia and failure to thrive were the most frequent presenting signs of GSD Ia, while hepatomegaly and recurrent bacterial infections (most frequently affecting skin) were the main clinical signs preceding diagnosis in GSD Ib.

In approximately 20% of children with GSD Ib neutropenia was not apparent at diagnosis. However, all of these patients developed neutropenia within three years after birth. Inflammatory bowel disease was verified in 20,6% patients with GSD Ib during follow up, with age at onset ranging from 2 to 10 years.

In siblings affected with GSD type Ib (Table S3), despite earlier establishment of diagnosis, the younger sibling failed to achieve normal height and has mild cognitive deficit. Hypogonadotropic hypogonadism present only in younger sibling may also reflect importance of metabolic control.

DISCUSSION

We analyzed one of the largest GSD Ib cohorts ever reported in a Caucasian population (7, 8). At the same time this is the largest cohort from Southeastern Europe and the first study which addresses the epidemiological, molecular and clinical characteristics of patients with GSD I from Serbian population.

Incidence of GSD I in Serbian population is among the highest in the world, surpassed only by the report from Saudi Arabia and projected incidence in Ashkenazi Jews calculated on the basis of mutation carrier's frequency (4, 9). Furthermore, predominance of GSD Ib among our patients represents a unique observation that has not been encountered in any other population so far. The epidemiological data from Serbian population demonstrate the highest worldwide incidence of GSD Ib which might be explained by founder effect and/or genetic drift.

In Caucasians, only p.Arg83Cys and p.Gln347Ter *G6PC* variants are repeatedly found (1). In our small cohort of GSD Ia patients, variant p.Gln347Ter was not found and p.Arg83Cys was highly predominant (90%).

Although majority of identified genetic variants in *SLC37A4* gene are considered private, recent reviews pointed out that p.Gly339Cys and p.Leu348Valfs*53 are the most common variants in Caucasian populations (1, 10). Interestingly, p.Gly339Cys variant was not detected in our study while p.Leu348Valfs*53 variant reached the highest ever reported frequency of 39% (7, 11). Even more frequent variant in Serbian GSD Ib patients is p.Asn27Lys (46.3%). Variants p.Asn27Lys and p.Ser54Arg were reported for the first time in patients from Southeastern Europe (7, 11) and since then have not been found in any other study.

Computational analyses predicted that novel c.785G>A variant could act as both splicing and missense variant. A wild-type *SLC37A4* gene has a weak acceptor splice at the beginning of exon 5 giving rise to a minor shorter transcript in some tissues besides a major full length transcript (6). This r.785_870del86 transcript found in traces in healthy controls and more abundantly in GSD patients leads to a truncated protein lacking the carboxyterminal sequence (-X-Lys-Lys-X-X) necessary for retention in the endoplasmic membrane (11). Our transcriptional analysis quantified that in heterozygous patients with c.785G>A variant, splicing efficacy of this acceptor splice site is further decreased by 60-80% in comparison to normal controls. Unlike intronic splice variants, an exonic variant will affect the protein even in the case of a leaky splicing. Since a splicing machinery succeeds in creating full length RNA, the resulting full length SLC37A4 protein is affected with a damaging amino acid substitution. Therefore, a dual nature of c.785G>A variant, resulting in either p.Ser263Glyfs*33 or p.Gly262Asp secures its pathogenic effect. In concordance, all four patients found in our cohort that carried c.785G>A had typical characteristics of GSD Ib (Table 2).

Although p.Gly135Asp has never been reported in a patient, it was functionally characterized as a part of functionally important “signature motif”. It was shown that p.Gly135Asp completely abolishes G6P uptake activity. Finally, we discovered the first patient with p.Gly135Asp (Table 2) further confirming the importance of the “signature motif” for SLC37A4 activity.

In addition to computational analyses which characterized novel p.Gly83Glu as damaging, findings that p.Leu85Pro and p.Gly88Asp have 0% and 2.3% of residual activity depict high importance of residues in the second helical trans-membrane domain on SLC37A4 activity (12). Moreover, two unrelated patients in whom p.Gly83Glu was found

presented characteristic GSD Ib symptoms including neutropenia from the early age (Table 2).

All patients from collected cohort have been successfully genotyped reaching genetic diagnostic rate of 100%. Presented spectrum of genetic variants which includes three newly described *SLC37A4* variants will facilitate medical genetic practice. We trust that epidemiological data presented herein should prompt higher awareness for GSD I in Serbia and surrounding countries of Southeastern Europe.

REFERENCES

1. Kishnani PS, Austin SL, Abdenur JE et al. Diagnosis and management of glycogen storage disease type I: a practice guideline of the American College of Medical Genetics and Genomics. Genetics in medicine : official journal of the American College of Medical Genetics 2014: 16: e1.
2. Vega AI, Medrano C, Navarrete R et al. Molecular diagnosis of glycogen storage disease and disorders with overlapping clinical symptoms by massive parallel sequencing. Genetics in medicine : official journal of the American College of Medical Genetics 2016: 18: 1037-1043.
3. Ozen H. Glycogen storage diseases: new perspectives. World journal of gastroenterology 2007: 13: 2541-2553.
4. Moammar H, Cheriyan G, Mathew R et al. Incidence and patterns of inborn errors of metabolism in the Eastern Province of Saudi Arabia, 1983-2008. Annals of Saudi medicine 2010: 30: 271-277.
5. Stojiljkovic M, Klaassen K, Djordjevic M et al. Molecular and phenotypic characteristics of seven novel mutations causing branched-chain organic acidurias. Clinical genetics 2016: 90: 252-257.
6. Ihara K, Nomura A, Hikino S et al. Quantitative analysis of glucose-6-phosphate translocase gene expression in various human tissues and haematopoietic progenitor cells. Journal of inherited metabolic disease 2000: 23: 583-592.
7. Janecke AR, Mayatepek E, Utermann G. Molecular genetics of type 1 glycogen storage disease. Molecular genetics and metabolism 2001: 73: 117-125.
8. Melis D, Fulceri R, Parenti G et al. Genotype/phenotype correlation in glycogen storage disease type 1b: a multicentre study and review of the literature. European journal of pediatrics 2005: 164: 501-508.
9. Ekstein J, Rubin BY, Anderson SL et al. Mutation frequencies for glycogen storage disease Ia in the Ashkenazi Jewish population. American journal of medical genetics Part A 2004: 129a: 162-164.
10. Froissart R, Piraud M, Boudjemline AM et al. Glucose-6-phosphatase deficiency. Orphanet journal of rare diseases 2011: 6: 27.
11. Santer R, Rischewski J, Block G et al. Molecular analysis in glycogen storage disease 1 non-A: DHPLC detection of the highly prevalent exon 8 mutations of the G6PT1 gene in German patients. Human mutation 2000: 16: 177.
12. Chou JY, Mansfield BC. The SLC37 family of sugar-phosphate/phosphate exchangers. Current topics in membranes 2014: 73: 357-382.

Figure Legends

Figure 1. Analyses of novel *SLC37A4* variants: (a) RT-PCR profile of healthy control (left) and patient with p.[Ser263Glyfs*33];[Asn27Lys] genotype (right), (b) wild-type and r.785_870del86 cDNA sequences (c) qRT-PCR study showing relative expression of r.785_870del86 *SLC37A4* transcript in mononuclear cells of patients in comparison to four healthy controls used as the calibrator, (d) Three-dimensional molecular models of SLC37A4 protein with close-up view of the regions harboring novel missense variants.

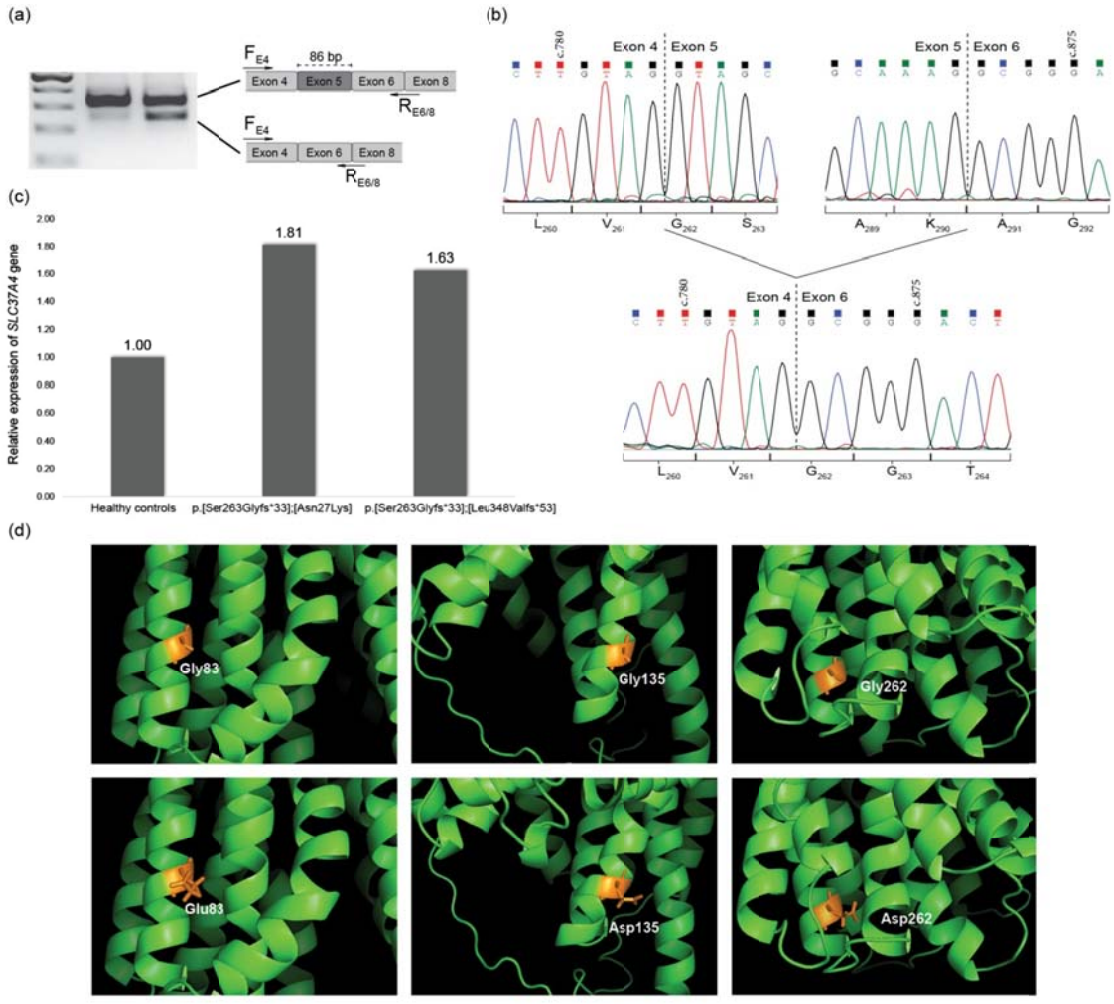

Table 1. Spectrum and frequency of genetic variants in unrelated Serbian patients with GSD I

GENE	Exon	Genetic variant		TYPE	Number of chromosomes	Relative frequency (%)	DESCRIPTION HGMD accession number is given between ()
		Nucleotide change*	Amino acid change				
G6PC	Exon 2	c.247C>T	p.Arg83Cys (p.R83C)	missense	9	90 %	Reported (CM930261)
	Exon 4	c.518T>C	p.Leu173Pro (p.L173P)	missense	1	10 %	Reported (CM077542)
SLC37A4	Exon 1	c.81T>A	p.Asn27Lys (p.N27K)	missense	25	46.30	Reported in a patient from Southeastern Europe (CM002021)
	Exon 2	c.162C>A	p.Ser54Arg (p.S54R)	missense	1	1.85	Reported in an Austrian/Slavic patient (CM002603)
	Exon 2	c.248G>A	p.Gly83Glu (p.G83E)	missense	2	3.70	Novel variant** PolyPhen-2 = 1, SIFT = 0 and MutPred= 0.907, ClustalW2 = absolute evolutionary conservation***
	Exon 3	c.404G>A	p.Gly135Asp (p.G135D)	missense	1	1.85	Novel variant PolyPhen-2 = 0.993, SIFT = 0 and MutPred = 0.962, ClustalW2 = absolute evolutionary conservation Functionally characterized (13)
	Exon 5	c.785G>A	p.Ser263Glyfs*33 or p.Gly262Asp (p.G262D)	Splice site/missense	4	7.40	Novel variant UMD-predictor score: wt - 72.65; mut - 69.52 Fruitfly score: wt - 0.78; mut - 0.60 PolyPhen-2 = 0.991, SIFT = 0 and MutPred =0.792, ClustalW2 = absolute evolutionary conservation
	Exon 8	c.1042_1043delCT	p.Leu348Valfs*53	frameshift deletion	21	38.90	Reported (CD982664)

*All identified nucleotide changes were numbered based on cDNA reference sequences and as recommended by the Human Genome Variation Society (http://www.hgvs.org/mutnomen). GenBank accession numbers for the sequences used in the analyses were as follows: NM_000151.3 (*G6PC*), NM_001164277.1 (*SLC37A4*).

**Genetic variants designated as novel were not found in the HGMD Professional database, Ensemble, EVS, ExAC or in TruSight One derived variant data base including 75 subjects from Serbian population, except c.785G>A which was found in ExAC on 1 allele among 58238 analyzed alleles.

***Computational analysis of the effect on splicing process was performed using different softwares: Berkeley Drosophila Genome Project BDGP (www.fruitfly.org/seq_tools/plice.html) and Human Splicing Finder (http://www.umd.be/HSF/). New missense genetic variants were analyzed with PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/), SIFT (http://sift.jcvi.org/) and MutPred (http://mutpred.mutdb.org). Amino acid alignments were constructed with ClustalW2 software (http://www.ebi.ac.uk/clustalw2/) using sequences of different vertebrate species (*zebrafish, chicken, mouse, gorilla and human*).

Table 2. Clinical characteristics of GSD Ib patients harboring novel genetic variants in *SLC37A4* gene

No.	Genotype		Clinical Vignette
P25	**p.Gly83Glu**	p.Leu348Valfs*53	The child was diagnosed with GSD type Ib at the age of 12 months with hepatomegaly as presenting sign. Throughout years, metabolic status featured elevated lactate 4,1-8,0 mmol/L and marked hyperlipidemia (average serum cholesterol of 9,1 mmol/L and serum triglycerides of 18 mmol/L). Final growth is below normal (definitive height 147 cm) and multiple hepatic adenoma were found at the age of 27 years with maximal diameter of 30 mm. Neutropenia is present from an early age but the frequency of serious bacterial infections remained low. No signs of inflammatory bowel disease are present in this patient.
P9	**p.Gly83Glu**	p.Leu348Valfs*53	Neutropenia was present since diagnosis (9 months of age) and at the age of 10 years myeloid status worsened significantly to chronic leukopenia (below 2 x 10(9)/L) and neutropenia (below 150 x 10(6)/L). Neupogen treatment did not result in significant elevation of neutrophil count but frequency of skin and respiratory infections reduced. Since the age of 10 years, the boy has intermittent flairs of diarrhea, bloating, abdominal pain and laboratory elements of protein losing enteropathy. Histopathology of colon mucosa has shown mild lymphoplasmocytic infiltration of mucosal lamina propria.
P22	**p.Gly135Asp**	p.Asn27Lys	At the age of 6 months "enlarged abdomen" was noticed, but diagnosis was established at 14 months on the basis of hepatomegaly and typical biochemical profile for GSD type I. Neutropenia was not present at the time of diagnosis, but it developed during second year of life.
P3	**p.Ser263Glyfs*33**	p.Leu348Valfs*53	There were four patients with combined heterozygosity that includes the novel variant c.785G>A. Three of those cases also harbor previously described c.81T>A variant, while one girl is heterozygous for the other well-known variant: c.1042_1043delCT. The presence of novel genetic variant c.785G>A was associated with an early onset of GSD signs and symptoms (1-6 months of age), most commonly presenting with hepatomegaly and bacterial infections. All patients developed neutropenia during infancy. One patient with c.81T>A/ c.785G>A genotype developed signs of inflammatory bowel disease at the age of 8 years.
P24	**p.Ser263Glyfs*33**	p.Asn27Lys	
P26	**p.Ser263Glyfs*33**	p.Asn27Lys	
P5	**p.Ser263Glyfs*33**	p.Asn27Lys	

New genetic variants were marked in bold; The segregation of mutated alleles was confirmed in both parents except in the patient P9 for whom only mother's sample was available and confirmed presence of p.L348Vfs*53.