



Genetic Profiles of Korean Patients With Glucose-6-Phosphate Dehydrogenase Deficiency

Jaewoong Lee, M.D.^{1,2}, Joonhong Park, Ph.D.^{1,2}, Hayoung Choi, B.S.², Jiyeon Kim, B.S.², Ahlm Kwon, B.S.², Woori Jang, Ph.D.^{1,2}, Hyojin Chae, Ph.D.^{1,2}, Myungshin Kim, Ph.D.^{1,2}, Yonggoo Kim, Ph.D.^{1,2}, Jae Wook Lee, Ph.D.³, Nack-Gyun Chung, Ph.D.³, and Bin Cho, Ph.D.³

Department of Laboratory Medicine¹, Catholic Genetic Laboratory Center², Department of Pediatrics³, College of Medicine, The Catholic University of Korea, Seoul, Korea

Background: We describe the genetic profiles of Korean patients with glucose-6-phosphate dehydrogenase (G6PD) deficiencies and the effects of *G6PD* mutations on protein stability and enzyme activity on the basis of *in silico* analysis.

Methods: In parallel with a genetic analysis, the pathogenicity of *G6PD* mutations detected in Korean patients was predicted *in silico*. The simulated effects of *G6PD* mutations were compared to the WHO classes based on G6PD enzyme activity. Four previously reported mutations and three newly diagnosed patients with missense mutations were estimated.

Results: One novel mutation (p.Cys385Gly, labeled G6PD Kangnam) and two known mutations [p.Ile220Met (G6PD São Paulo) and p.Glu416Lys (G6PD Tokyo)] were identified in this study. *G6PD* mutations identified in Koreans were also found in Brazil (G6PD São Paulo), Poland (G6PD Seoul), United States of America (G6PD Riley), Mexico (G6PD Guadalajara), and Japan (G6PD Tokyo). Several mutations occurred at the same nucleotide, but resulted in different amino acid residue changes in different ethnic populations (p.Ile380 variant, G6PD Calvo Mackenna; p.Cys385 variants, Tomah, Madrid, Lynwood; p.Arg387 variant, Beverly Hills; p.Pro396 variant, Bari; and p.Pro396Ala in India). On the basis of the *in silico* analysis, Class I or II mutations were predicted to be highly deleterious, and the effects of one Class IV mutation were equivocal.

Conclusions: The genetic profiles of Korean individuals with *G6PD* mutations indicated that the same mutations may have arisen by independent mutational events, and were not derived from shared ancestral mutations. The *in silico* analysis provided insight into the role of *G6PD* mutations in enzyme function and stability.

Key Words: Glucose-6-phosphate dehydrogenase deficiency, Mutation, *In silico* analysis, Korean

Received: April 7, 2016

Revision received: June 24, 2016

Accepted: November 29, 2016

Corresponding author: Joonhong Park
Department of Laboratory Medicine,
College of Medicine, The Catholic
University of Korea, Daejeon St. Mary's
Hospital, 64 Daeheung-ro, Jung-gu,
Daejeon 34943, Korea
Tel: +82-42-220-9799
Fax: +82-42-220-9915
E-mail: miziro@catholic.ac.kr

Co-corresponding author: Bin Cho
Department of Pediatrics, College of
Medicine, The Catholic University of Korea,
Seoul St. Mary's Hospital, 222 Banpo-
daero, Seocho-gu, Seoul 06591, Korea
Tel: +82-2-2258-6187
Fax: +82-2-537-4544
E-mail: chobinkr@catholic.ac.kr

© Korean Society for Laboratory Medicine

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

Glucose-6-phosphate dehydrogenase (G6PD) deficiency is the most prevalent X-linked enzymopathy. G6PD is the first enzyme in the pentose phosphate pathway, and NADPH generated by the pathway provides an important source for intracellular reduction, particularly for red blood cells (RBCs) [1]. Since G6PD

is the only NADPH-producing enzyme in RBCs, its activity in these cells provides defense against oxidative damage. Acute hemolytic anemia is a common clinical symptom of the deficiency, but G6PD-deficient individuals usually have no clinical manifestations and remain asymptomatic until they are exposed to a hemolytic trigger. The triggers include various exogenous agents, such as infection and hemolysis-inducing drugs, and

can each cause jaundice, hyperbilirubinemia, and hemoglobinuria. When a G6PD deficiency is suspected, a patient receives various tests, including a complete blood count (CBC) with reticulocyte count, direct and indirect bilirubin levels, lactate dehydrogenase (LDH), Coombs test, and G6PD enzyme activity. A genetic analysis by *G6PD* sequencing is also available.

According to the WHO classification, G6PD deficiency is divided into five classes on the basis of the severity of the enzyme deficiency as measured by the level of RBC G6PD activity and clinical manifestations [2]. The majority of patients with G6PD deficiency belong to Class II, characterized by a severe enzyme deficiency, but rare G6PD-deficient individuals fall into Class I, with an even more severe enzyme deficiency related to chronic non-spherocytic hemolytic anemia (CNSHA). Genetic diagnostic methods can be used to identify asymptomatic patients who are not in an acute aggravation state, even those with a Class IV G6PD deficiency, with enzyme activity levels within the normal, reference range, but who have the potential for aggravation in response to triggers.

Since G6PD Riley and Guadalajara were first reported by our institute [3, 4], two additional G6PD deficiency patients have been genetically confirmed in Korea [5, 6]. We described three more Korean cases of genetically confirmed G6PD deficiency, covering the laboratory profiles of all seven patients including previously reported cases, and investigated mutations in *G6PD* using an *in silico* approach. We also compared the simulated effects of the *G6PD* mutations to WHO classes according to the level of enzyme activity in RBCs and clinical manifestations.

METHODS

1. Patients

All seven known Korean male patients with *G6PD* mutations including four previously reported cases were examined. The seven patients experienced episodes of acute aggravation of hemolytic anemia with decreased G6PD enzyme activity. Among them, three patients were newly diagnosed as G6PD-deficient in this study. The G6PD enzyme activity levels in the RBCs of all three patients were low, i.e., 10.5, 2.1, and 0.8, respectively (reference range for men: 7.9–16.3 U/g Hb). The study protocol was approved by the Institutional Review Board of The Catholic University of Korea, and written informed consent for clinical and molecular analyses was obtained from the three newly diagnosed cases.

2. Biochemical analysis of G6PD enzyme activity levels

A spectrophotometric assay was used to quantify G6PD enzyme activity (Ben S.r.l. Biochemical Enterprise, Milan, Italy) by measuring the formation of NADPH molecules (based on absorbance at 340 nm). Fluorescence was detected by using a Hitachi U-3010 UV-Visible, Scanning Spectrophotometer (Hitachi, Tokyo, Japan).

3. Direct sequencing

A genetic analysis was performed by direct sequencing of the *G6PD*. Genomic DNA was extracted from peripheral blood by using a QIAamp DNA Mini Kit (Qiagen GmbH, Hilden, Germany). Entire coding exons and flanking intronic sequences of *G6PD* were amplified by PCR using different combinations of 11 primer sets designed using Primer3 (<http://bioinfo.ut.ee/primer3/>) by the authors. Direct sequencing of PCR products was performed by using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA), and the products were resolved on the ABI 3130XL Genetic Analyzer (Applied Biosystems). Sequence electropherograms were analyzed by using Sequencher 4.9 (Gene Codes, Ann Arbor, MI, USA). The *G6PD* sequence with RefSeq ID NM_001042351.2 was used as a reference for cDNA nucleotide numbering. All identified variants were confirmed by bidirectional resequencing.

4. *In silico* analysis of identified amino acid residues

1) Sequence variant databases

Other *G6PD* mutations at the identified amino acid residues were retrieved from the Human Gene Mutation Database (HGMD, <http://www.hgmd.cf.ac.uk/>) [7], ClinVar (<http://www.ncbi.nlm.nih.gov/clinvar/>) [8], and the 1000 Genomes Project (<http://browser.1000genomes.org/>) [9]. The Exome Aggregation Consortium (ExAC, <http://exac.broadinstitute.org/>) was also searched; this source provides exome sequencing data from a variety of large-scale sequencing projects, including 60,706 unrelated individuals, and serves as a useful reference set of allele frequencies for severe disease studies. An allele frequency of less than 0.001 indicates that a variant is rare in the healthy population and the variant can be interpreted as a pathogenic mutation. The Exome Variant Server (EVS, <http://evs.gs.washington.edu/EVS/>), which includes 6,503 samples drawn from multiple ESP (Exome Sequencing Project) cohorts, representing all ESP exome variant data, the Korean Reference Genome DataBase (KRGDB, <http://152.99.75.168/KRGDB/menuPages/firstInfo.jsp/>), which harbors whole genome sequencing data for 622 Korean individuals, and the Total Integrated Archive of short-Read and Array (TIARA, <http://tiara.gmi.ac.kr/>) database, including data gener-

ated by next-generation sequencing and high-resolution comparative genomic hybridization arrays, were also searched.

2) Computational prediction tools

There are many prediction tools for the effects of genetic variants that use various methods for prediction. These tools provide useful information, but sometimes the predictions for the same variant do not agree. Thus, several prediction tools were compared to evaluate the mutations identified in Korean patients. On the basis of this comparative analysis, the most suitable method for predicting the effects of *G6PD* mutations was identified. The following widely established computational prediction methods were applied: the evolution-based prediction tools including Sorting Intolerant From Tolerant (SIFT, <http://sift.jcvi.org/>) [10], the Protein Variation Effect Analyzer (PROVEAN, <http://provean.jcvi.org/>) [11], Multivariate Analysis of Protein Polymorphisms (MAPP, <http://mendel.stanford.edu/SidowLab/downloads/MAPP/>) [12], and Align-GVGD (http://agvgd.iarc.fr/agvgd_input.php) [13]. Protein structure and function were assessed by using straightforward comparative physical and evolutionary analyses and sequence- and structure-based bioinformatics tools, such as Polymorphism Phenotyping v2 (PolyPhen-2, <http://genetics.bwh.harvard.edu/pph2/>) [14], MutPred (<http://mutpred.mutdb.org/>) [15], and SnpEff 4.0 (<http://snpeff.switchlab.org/>) [16].

3) Evolutionary conservation analysis

To assess the extent of conservation of the *G6PD* amino acid residues thought to be associated with disease, the deduced amino acid sequences were compared by aligning the protein sequences of various vertebrate species obtained from the Evolutionary Annotation Database (Evola, <http://www.h-invitational.jp/evola/>) and conservation was then quantified by using Scorecons [17]. The Scorecons analysis generates a composite score that accounts for amino acid frequencies, stereochemical diversity, gap penalties, and sequence weighting (http://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/valdar/scorecons_server.pl). The overall diversity of all position scores for the *G6PD* protein sequence was 0.529, ranging from 0 for non-conserved residues to 1 for highly conserved residues. A *G6PD* multiple sequence alignment was constructed using the following taxa, which represent broad phylogenetic diversity: human (*Homo sapiens*; HIT-000029531), orangutan (*Pongo*; ENSPPYT00000024358), macaque (*Macaca*; XM_001095273), mouse (*Mus*; BC075663), rat (*Rattus*; BC081820), horse (*Equus*; XM_001492232), opossum (*Monodelphis*; XM_001362790), zebrafish (*Danio*; XM_694076), medaka (*Oryzias*; ENSORLT00000015543), and fugu (*Takifugu*; SINFRUT00000143893).

jp/evola/) and conservation was then quantified by using Scorecons [17]. The Scorecons analysis generates a composite score that accounts for amino acid frequencies, stereochemical diversity, gap penalties, and sequence weighting (http://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/valdar/scorecons_server.pl). The overall diversity of all position scores for the *G6PD* protein sequence was 0.529, ranging from 0 for non-conserved residues to 1 for highly conserved residues. A *G6PD* multiple sequence alignment was constructed using the following taxa, which represent broad phylogenetic diversity: human (*Homo sapiens*; HIT-000029531), orangutan (*Pongo*; ENSPPYT00000024358), macaque (*Macaca*; XM_001095273), mouse (*Mus*; BC075663), rat (*Rattus*; BC081820), horse (*Equus*; XM_001492232), opossum (*Monodelphis*; XM_001362790), zebrafish (*Danio*; XM_694076), medaka (*Oryzias*; ENSORLT00000015543), and fugu (*Takifugu*; SINFRUT00000143893).

RESULTS

1. Genetic and phenotypic characteristics of *G6PD* deficiencies in Korean patients

Seven missense mutations were identified as hemizygous (Table 1) in male patients with *G6PD* deficiencies. Among the seven mutations, five were identified at our institute; one novel mutation (p.Cys385Gly named as *G6PD* Kangnam) and two known mutations (p.Ile220Met [*G6PD* São Paulo] [18] and p.Glu416Lys [*G6PD* Tokyo] [19]) were newly identified in the Korean population in this study. Two known mutations (p.Arg387Cys [*G6PD* Guadalajara] [20] and p.Ile380Thr [*G6PD* Riley] [21]) were reported previously by other researchers [3, 4] at our institute. In addition to these five mutations, we found two missense mutations (p.Gly306Ser [*G6PD* Seoul] [5] and p.Pro396Arg [unnamed]

Table 1. Missense mutations in the *G6PD* reported in Korean patients with *G6PD* deficiency

Case No.	S/A	G6PD (U/g Hb)	WHO category	Hb (g/L)	MCV (fL)	MCH (pg)	MCHC (%)	Reti (%)	LDH (U/L)	T-Bil (mg/dL)	cDNA change	Protein change	Location	Annotation	References
1	M/7	1.8	Class I	0.26	NA	NA	NA	NA	NA	NA	c.1139T>C	p.Ile380Thr	Exon10	Riley	Park <i>et al</i> , 1999 [3]
2	NA/NA	NA	Class II	NA	NA	NA	NA	NA	NA	NA	c.916G>A	p.Gly306Ser	Exon9	Seoul	Beutler <i>et al</i> , 2002 [5]
3	M/2	0.5	Class I	1.21	10.5	33.3	31.2	10.4	1,329	5.56	c.1159C>T	p.Arg387Cys	Exon10	Guadalajara	Lee <i>et al</i> , 2004 [4]
4	M/2	0.2	Class I	1.02	NA	NA	NA	15.6	NA	NA	c.1187C>G	p.Pro396Arg	Exon10	No name	Jang <i>et al</i> , 2015 [6]
5	M/3	10.5	Class IV	1.21	97.0	33.3	34.4	9.65	2,000	14.83	c.660C>G	p.Ile220Met	Exon7	São Paulo	This study
6	M/5	2.1	Class I	1.11	95.1	31.9	33.5	3.63	1,151	1.1	c.1153T>G	p.Cys385Gly	Exon10	Kangnam*	This study
7	M/2	0.8	Class I	0.69	94.9	31.2	32.8	5.82	459	0.6	c.1246G>A	p.Glu416Lys	Exon10	Tokyo	This study

*Novel mutation identified and annotated in this study.

Abbreviations: *G6PD*, glucose-6-phosphate dehydrogenase; S/A, sex/age (years); MCH, mean corpuscular hemoglobin; MCV, mean corpuscular volume; MCHC, mean corpuscular hemoglobin concentration; Reti, reticulocyte; T-Bil, total bilirubin; LDH, lactate dehydrogenase; NA, not available.

[6]) in a literature review. Mutations affecting the same residue (Cys385) as that of G6PD Kangnam, but resulting in different amino acid changes, p.Cys385Arg (G6PD Tomah) [22], p.Cys385Trp (G6PD Madrid) [23], and p.Cys385Phe (Lynwood) [24], have been described in different ethnic groups.

The three newly diagnosed patients as having G6PD deficiency were characterized. Patient of case 5 was a 3-yr-old boy who suffered from fever and jaundice. He had no family history of hematologic disorders. CBC showed the following: Hb 0.96 g/L, mean corpuscular volume (MCV) 101.3 fL, mean corpuscular Hb (MCH) 32.3 pg, mean corpuscular Hb concentration (MCHC) 31.9%, reticulocyte 17.18%, and undetectable G6PD enzyme activity. The Coombs test and osmotic fragility test were negative. After one week, he recovered from the acute aggravation. The G6PD enzyme level increased to 10.5 U/g Hb, and the G6PD deficiency was categorized as Class IV according to the WHO classification. Sanger sequencing of the *G6PD* revealed a c.660C>G (p.Ile220Met) mutation, which was previously reported as *G6PD São Paulo* [18].

Patient of case 6 was a 5-yr-old boy who suffered from a fever and a pale appearance. He had no family history of hematologic disorders. He visited our hospital after recovery from the acute aggravation. The Coombs test and the osmotic fragility test were negative. The G6PD enzyme level was 2.1 U/g Hb without acute aggravation. His condition was classified as Class I G6PD deficiency according to the WHO classification owing to the severely decreased enzyme activity with CNSHA. Sanger sequencing revealed a c.1153T>G (p.Cys385Gly) mutation in *G6PD*. This mutation was not reported previously; therefore, we named the mutation *G6PD Kangnam*. Interestingly, mutations affecting the same residue as that of *G6PD Kangnam*, but resulting in different amino acid changes, p.Cys385Arg (*G6PD Tomah*) [22], p.Cys385Trp (*G6PD Madrid*) [23], and p.Cys385Phe (*G6PD Lynwood*) [24], have been previously described in different ethnic groups.

Patient of case 7 was a 2-yr-old boy who had a fever and a pale appearance. He had a family history of anemia (maternal uncle and younger brother). With acute aggravation, the G6PD level was 0.8 U/g Hb. His condition was classified as Class I G6PD deficiency according to the WHO classification. Sanger sequencing revealed a c.1246G>A (p.Glu416Lys) mutation, which is known as *G6PD Tokyo* [19].

Four additional mutations were previously reported in Korean individuals. A case study of patient 2, with the *G6PD Seoul* mutation, was not published, but was described in a later study [5]. The mutation was c.916G>A in exon 9 and a patient with the

same mutation was classified as Class II in a later report (the G6PD activity level was 1.6 U/g Hb and there was no CNSHA). Patient 1 was a 7-yr-old boy who presented with mild jaundice and dyspnea. He had neonatal jaundice and recurrent nonspherocytic hemolytic anemia with type B hepatitis. CBC showed an Hb of 0.26 g/L and G6PD activity of 1.8 U/g Hb. On the basis of CNSHA, the patient was Class I according to the WHO classification. He had three brothers who died from severe neonatal jaundice and acute hemolytic anemia. The mutation was c.1139T>C on exon 10 of *G6PD*. Patient 3 was a 22-month-old boy who presented with jaundice and anemia. He had no family history of hematologic disorders. CBC results were as follows: Hb 0.62 g/L, MCH 33.3 pg, MCHC 31.2%, MCV 106.6 fL and reticulocyte 10.4%. A peripheral blood smear revealed anisocytosis, poikilocytosis, and macrocytosis. Total bilirubin was 5.56 mg/dL, direct bilirubin was 0.57 mg/dL, LDH was 1,329 U/L, and G6PD activity was 0.5 U/g Hb. The patient belonged to Class I according to the WHO classification. Based on sequencing, the mutation was c.1159C>T in exon 10 of *G6PD*. Patient 4 was a 20-month-old boy who presented with chronic hemolytic anemia and several episodes of acute exacerbation. His Hb level was 1.02 g/L, and G6PD enzyme level was 0.2 U/g Hb. The patient belonged to Class I according to the WHO classification. The mutation was documented as c.1187C>G in exon 10.

2. Effect of mutations on disease manifestation

In an *in silico* analysis, *G6PD* mutations associated with Class I or II were predicted to be highly deleterious on the basis of their effects on protein structure and function, whereas the effects of Class IV mutations were unclear. Scorecons, an evolutionary conservation predictor, showed that the residues were located in highly conserved regions (Table 2). In terms of evolution-based sequence diversity, the overall diversity based on all position scores for the *G6PD* protein sequence was 0.529, ranging from 0 for non-conserved residues to 1 for highly conserved residues. Mutations affecting the same residue, such as p.Ile380, p.Cys385, p.Arg387, and p.Pro396, were predicted to be pathogenic on the basis of codon changes (Table 3). FoldX predicted alterations in protein stability, except for *G6PD São Paulo*, which has been reported in the ExAC database at an extremely low frequency, i.e., 1.141×10^{-5} (Fig. 1A and B, Table 4). For the Class IV mutation, *G6PD São Paulo*, an *in silico* analysis using SIFT, PROVEAN, PolyPhen-2, Align-GVGD, and the FoldX implemented in SNPeff 4.0 predicted a relatively mild effect, consistent with the observed enzyme activity levels.

Table 2. *In-silico* analysis of missense mutations of the *G6PD* gene reported in Korean patients with G6PD deficiency

Annotation	G6PD (U/g Hb)	WHO category	SIFT	PROVEAN	PolyPhen-2	MutPred	MAPP	Align-GVGD*	Scorecons [†]	FoldX (Stability)
São Paulo	10.5	Class IV	Tolerated (0.19)	Neutral (-2.315)	Possibly damaging (0.482)	Deleterious (0.638)	Bad (15.70)	Class C0 (GD 10.12)	Conserved (1)	No effect
Seoul	NA	Class II	Damaging (0.02)	Deleterious (-5.521)	Probably damaging (0.975)	Deleterious (0.879)	Good (7.5)	Class C55 (GD 55.27)	Conserved (1)	Reduced
Riley	1.8	Class I	Damaging (0.00)	Deleterious (-4.355)	Probably damaging (0.998)	Deleterious (0.886)	Bad (19.6)	Class C65 (GD 89.28)	Conserved (1)	Slightly reduced
Kangnam*	2.1	Class I	Tolerated (0.20)	Deleterious (-4.236)	Benign (0.007)	Deleterious (0.750)	Bad (15.41)	Class C65 (GD 158.23)	Conserved (1)	Slightly reduced
Guadalajara	0.5	Class I	Damaging (0.00)	Deleterious (-7.069)	Probably damaging (0.998)	Deleterious (0.857)	Bad (29.01)	Class C65 (GD 179.53)	Conserved (1)	Reduced
No name	0.2	Class I	Damaging (0.00)	Deleterious (-8.120)	Probably damaging (1.000)	Deleterious (0.915)	Bad (41.48)	Class C65 (GD 102.71)	Conserved (1)	Reduced
Tokyo	0.8	Class I	Tolerated (0.30)	Deleterious (-2.574)	Probably damaging (0.991)	Deleterious (0.722)	Bad (19.98)	Class C55 (GD 56.87)	Conserved (1)	Slightly enhanced

*Novel mutation identified and annotated in this study; [†]Multiple alignments of different vertebrate species obtained from Evolutionary annotation database (<http://www.h-invitational.jp/evola/>).

Abbreviations: G6PD, glucose-6-phosphate dehydrogenase; NA, not available.

Table 3. *In-silico* analysis of missense mutations of *G6PD* located on same position of amino acid residues identified in Korean G6PD deficiencies

Annotation	Protein change	WHO category	SIFT	PROVEAN	PolyPhen-2	MutPred	MAPP	Align GVGD*	Scorecons*
Riley (Case 3)	p.Ile380Thr	Class I	Damaging (0.00)	Deleterious (-4.355)	Probably damaging (0.998)	Deleterious (0.886)	Bad (19.6)	Class C65 (GD 89.28)	Conserved (1)
Calvo Mackenna	p.Ile380Val	Class I	Tolerated (0.16)	Neutral (-0.904)	Possibly damaging (0.705)	Deleterious (0.744)	Bad (10.65)	Class C25 (GD 28.68)	Conserved (1)
Kangnam [†] (Case 4)	p.Cys385Gly	Class I	Tolerated (0.20)	Deleterious (-4.236)	Benign (0.007)	Deleterious (0.750)	Bad (15.41)	Class C65 (GD 158.23)	Conserved (1)
Tomah	p.Cys385Arg	Class I	Tolerated (0.21)	Deleterious (-4.100)	Probably damaging (0.976)	Deleterious (0.829)	Bad (34.86)	Class C65 (GD 179.53)	Conserved (1)
Madrid	p.Cys385Trp	Class I	Tolerated (0.11)	Deleterious (-4.446)	Probably damaging (0.999)	Deleterious (0.774)	Bad (32.95)	Class C65 (GD 214.36)	Conserved (1)
Lynwood	p.Cys385Phe	Class I	Tolerated (0.68)	Deleterious (-3.801)	Probably damaging (0.976)	Deleterious (0.724)	Bad (21.32)	Class C65 (GD 204.39)	Conserved (1)
Guadalajara (Case 5)	p.Arg387Cys	Class I	Damaging (0.00)	Deleterious (-7.069)	Probably damaging (0.998)	Deleterious (0.857)	Bad (29.01)	Class C65 (GD 179.53)	Conserved (1)
Beverly Hills	p.Arg387His	Class I	Damaging (0.01)	Deleterious (-4.456)	Probably damaging (1.000)	Deleterious (0.834)	Bad (14.41)	Class C25 (GD 28.82)	Conserved (1)
Not annotated (Case 6)	p.Pro396Arg	Class I	Damaging (0.00)	Deleterious (-8.120)	Probably damaging (1.000)	Deleterious (0.915)	Bad (41.48)	Class C65 (GD 102.71)	Conserved (1)
Not annotated	p.Pro396Ala	Class I	Damaging (0.00)	Deleterious (-7.218)	Probably damaging (0.997)	Deleterious (0.921)	Bad (20.00)	Class C25 (GD 26.87)	Conserved (1)
Bari	p.Pro396Leu	Class I	Damaging (0.00)	Deleterious (-9.023)	Probably damaging (1.000)	Deleterious (0.909)	Bad (26.92)	Class C65 (GD 97.78)	Conserved (1)

*Multiple alignments of different vertebrate species obtained from Evolutionary annotation database (<http://www.h-invitational.jp/evola/>); [†]Novel mutation identified and annotated in this study.

Abbreviation: G6PD, glucose-6-phosphate dehydrogenase.

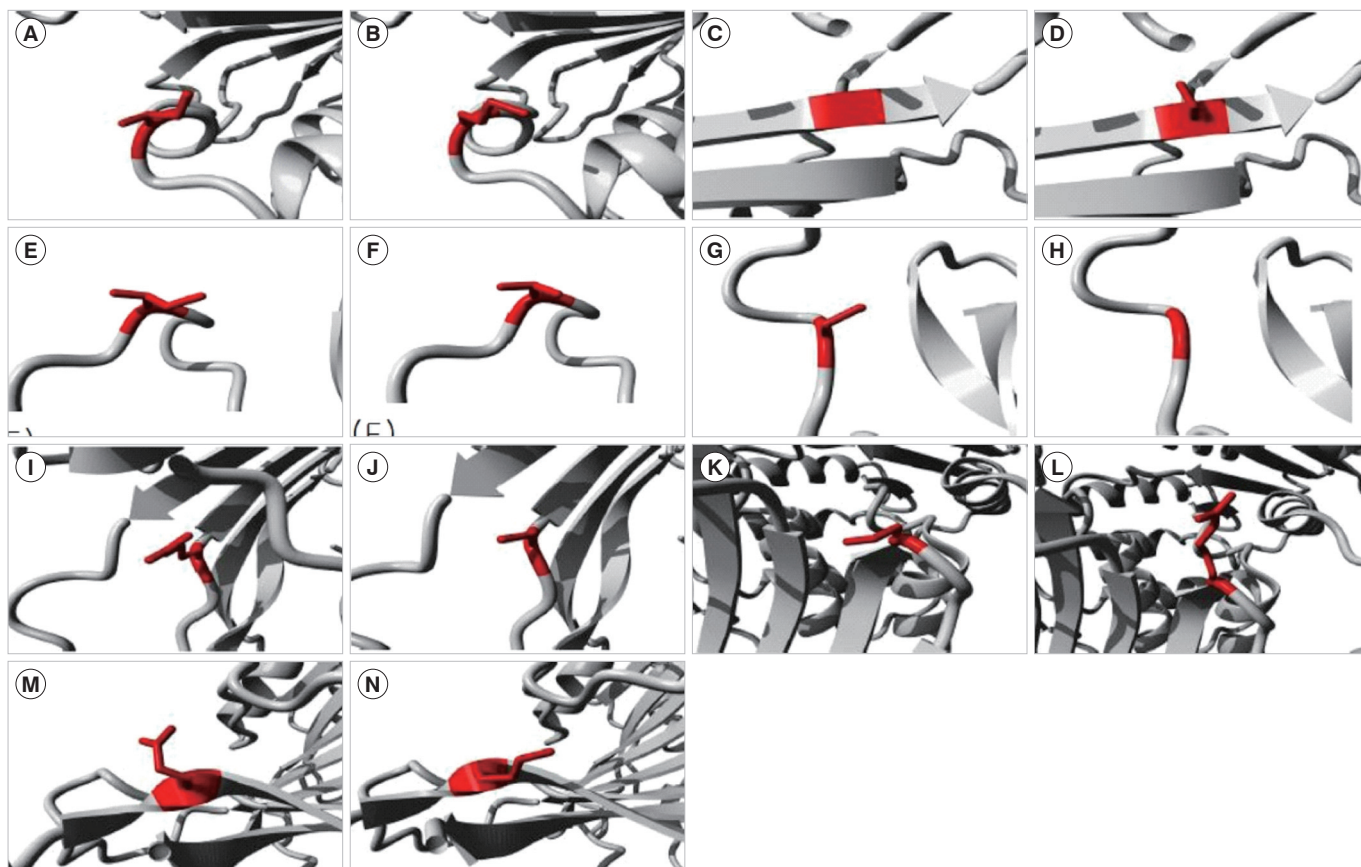


Fig. 1. Molecular visualization of the wild-type (WT) (left) and variant (right) amino acid using the FoldX function of SNPeff 4.0. If the mutation destabilizes the structure, $\Delta\Delta G$ is increased, whereas stabilizing mutations decrease $\Delta\Delta G$. Since the FoldX error margin is approximately 0.5 kcal/mol, changes in this range are considered insignificant. (A, B) The mutation from Ile (red in A) to Met (red in B) at position 220 resulted in a $\Delta\Delta G$ of -0.38 kcal/mol. This implies that the mutation had no effect on protein stability. (C, D) The mutation from Gly (red in C) to Ser (red in D) at position 306 resulted in a $\Delta\Delta G$ of 2.81 kcal/mol. This implies that the mutation reduced protein stability. (E, F) The mutation from Ile (red in E) to Thr (red in F) at position 380 resulted in a $\Delta\Delta G$ of 0.56 kcal/mol. This implies that the mutation slightly reduced protein stability. (G, H) The mutation from Cys (red in G) to Gly (red in H) at position 385 resulted in a $\Delta\Delta G$ of 0.65 kcal/mol. This implies that the mutation slightly reduced protein stability. (I, J) The mutation from Arg (red in I) to Cys (red in J) at position 387 resulted in a $\Delta\Delta G$ of 2.18 kcal/mol. This implies that the mutation reduced protein stability. (K, L) The mutation from Pro (red in K) to Arg (red in L) at position 396 resulted in a $\Delta\Delta G$ of 3.19 kcal/mol. This implies that the mutation reduced protein stability. (M and N) The mutation from Glu (red in M) to Lys (red in N) at position 416 resulted in a $\Delta\Delta G$ of -0.86 kcal/mol. This implies that the mutation slightly enhanced protein stability.

DISCUSSION

G6PD deficiencies selectively affect RBCs via two mechanisms. First, most known mutations decrease enzyme stability. Since these cells do not have the ability to synthesize proteins and replenish their enzyme levels, the enzyme level decreases as cells age during their 120-day lifespan in circulation. Second, G6PD activity is decreased, and the diminished ability of RBCs to withstand stress increases the risk of destruction by hemolysis.

To date, approximately 186 G6PD mutations have been documented, most of which (85%, 159/186) are single nucleotide substitutions leading to missense variants [25]. This and previ-

ous studies have revealed identical G6PD mutations in different countries around the world. Since 1999, when G6PD Riley was reported at our institute, seven G6PD mutation types have been identified in Korea. The same mutations have also been found in Brazil (G6PD São Paulo) [18], Poland (G6PD Seoul) [26], Americas (G6PD Riley) [21], Mexico (G6PD Guadalajara) [20], and Japan (G6PD Tokyo) [19]. In addition, several mutations can occur at the same nucleotide position, but result in different amino acid changes in different ethnic populations. For example, p.Cys385 variants have been reported as G6PD Kangnam as well as G6PD Tomah [22], Madrid [23], and Lynwood [24]. In addition, p.Pro396 variants have been described as p.Pro396Arg in Korean individ-

Table 4. Public variant databases of missense mutations of *G6PD* located on same position of amino acid residues identified in Korean G6PD deficiencies

	HGMD	ClinVar	ExAc	EVS	KRGDB	TIARA
Ethnicity			East and South Asian, African, European, Latino, and other	European- and African-American	Korean	Korean
Samples			43,806	6,503	622	28*
X:153762360 (p.Ile220Met)	CM0910658	Not reported	1.141 × 10 ⁻⁵ (2.087 × 10 ⁻⁵ in European)	Not reported	Not reported	Not reported
X:153761292 (p.Gly306Ser)	CM952139	Not reported	Not reported	Not reported	Not reported	Not reported
X:153760916 (p.Cys385Gly)	Not reported	Not reported	Not reported	Not reported	Not reported	Not reported
X:153760916 (p.Cys385Arg)	CM930277	rs137852322	Not reported	Not reported	Not reported	Not reported
X:153760914 (p.Cys385Trp)	CM970548	Not reported	Not reported	Not reported	Not reported	Not reported
X:153760915 (p.Cys385Phe)	CM973148	Not reported	Not reported	Not reported	Not reported	Not reported
X:153760930 (p.Ile380Thr)	CM950508	Not reported	Not reported	Not reported	Not reported	Not reported
X:153760931 (p.Ile380Val)	CM950509	Not reported	Not reported	Not reported	Not reported	Not reported
X:153760910 (p.Arg387Cys)	CM920285	rs137852334	Not reported	Not reported	Not reported	Not reported
X:153760909 (p.Arg387His)	CM930279	rs137852321	Not reported	Not reported	Not reported	Not reported
X:153760882 (p.Pro396Arg)	Not reported	Not reported	Not reported	Not reported	Not reported	Not reported
X:153760882 (p.Pro396Ala)	CM067413	Not reported	Not reported	Not reported	Not reported	Not reported
X:153760883 (p.Pro396Leu)	CM940798	Not reported	Not reported	Not reported	Not reported	Not reported
X:153760823 (p.Glu416Lys)	CM920288	Not reported	Not reported	Not reported	Not reported	Not reported

HGMD, Human Gene Mutation Database, <http://www.hgmd.cf.ac.uk/>.

ClinVar, <http://www.ncbi.nlm.nih.gov/clinvar/>.

ExAc, The Exome Aggregation Consortium, <http://exac.broadinstitute.org/>.

EVS, Exome Variant Server, <http://evs.gs.washington.edu/>.

KRGDB, Korean Reference Genome DB, <http://152.99.75.168/KRGDB/>.

TIARA, Totally Integrated high-resolution Array cgh data and high-throughput sequencing Read data Archive, <http://tiara.gmi.ac.kr/>.

*12 whole genomes and 16 transcriptomes

Abbreviation: G6PD, glucose-6-phosphate dehydrogenase.

uals [6], p.Pro396Ala in Indian individuals [27], and p.Pro396Leu in Italian individuals [28]. These findings and previous studies imply that the same mutations may have arisen by independent mutation events in Korea, rather than being inherited from common ancestral mutations [18-21, 26].

All of the *G6PD* mutations affected coding regions, and none has been described in a regulatory region. This suggests that reduced enzyme activity levels are associated with decreased enzyme activity in the G6PD-deficient phenotype, rather than deficiencies in gene expression [25, 29]. With respect to genotype-phenotype associations, the frequencies of Class I mutations found in exon 10 are significantly higher than those in other exons in G6PD deficiency cases [25]. Mutations located in this region, which encodes the binding interface between the subunits, have a highly deleterious effect on enzyme activity by disrupting the quaternary structure and stability of the protein, and appear to be the most easily identifiable in the general population [1].

In parallel with a genetic analysis, the pathogenicity of each *G6PD* mutation detected in the Korean patients was predicted by using an *in silico* analysis according to WHO classes. Class I mutations were predicted to be highly deleterious, while the effects of one Class IV mutation were equivocal. However, all mutations in Korean patients with G6PD deficiencies were located in residues that were conserved across various species. A high degree of evolutionary conservation of certain gene regions is essential for enzyme function and cell survival. Similar to a previous study [18], our G6PD São Paulo mutation with G6PD activity of 10.5 U/g Hb was categorized as Class IV, near the dimer interface. Rare Class IV mutations [5, 18, 30-32] showing nearly normal G6PD activity can be overlooked because molecular analyses of G6PD are only considered when episodes of hemolytic crisis occur.

In summary, identical *G6PD* mutations associated with G6PD deficiencies in Korean patients may have arisen by independent mutation events. An *in silico* analysis provided insight into the

roles of *G6PD* mutations from evolutionary and structure-based computational points of view. Further investigations involving more Koreans with G6PD deficiencies are needed to clarify the relationship between the clinical manifestations and mutational spectrum related to this enzyme deficiency.

Authors' Disclosures of Potential Conflicts of Interest

No potential conflicts of interest relevant to this article were reported.

Acknowledgments

We are grateful to The Catholic Genetic Laboratory Center for assisting us in this study and compiling this report. This study was supported by a grant from the Korea Health Technology R&D Project, Ministry of Health & Welfare, Republic of Korea (A120175).

REFERENCES

- Cappellini MD and Fiorelli G. Glucose-6-phosphate dehydrogenase deficiency. *Lancet* 2008;371:64-74.
- Glucose-6-phosphate dehydrogenase deficiency. WHO Working Group. *Bull World Health Organ* 1989;67:601-11.
- Park SW, Lee HJ, Lee W, Whang KT. A Case of Glucose-6-phosphate dehydrogenase Deficiency Causing Hemolytic Anemia. *Korean J Hematol* 1999; 34:334-7.
- Lee GB, Lee SJ, Kim YJ, Kim SY, Kim HH, Cho B, et al. A Case of G-6-PD Guadalajara. *Korean J Pediatr* 2004;47:210-3.
- Beutler E and Vulliamy TJ. Hematologically important mutations: glucose-6-phosphate dehydrogenase. *Blood Cells Mol Dis* 2002;28:93-103.
- Jang M-A, Kim J-Y, Lee K-O, Kim S-H, Koo HH, Kim H-J. A Novel de novo Mutation in the G6PD Gene in a Korean Boy with Glucose-6-phosphate Dehydrogenase Deficiency: Case Report. *Ann Clin Lab Sci* 2015; 45:446-8.
- Stenson P, Mort M, Ball EV, Shaw K, Phillips A, Cooper DN. The Human Gene Mutation Database: building a comprehensive mutation repository for clinical and molecular genetics, diagnostic testing and personalized genomic medicine. *Hum Genet* 2014;133:1-9.
- Landrum MJ, Lee JM, Riley GR, Jang W, Rubinstein WS, Church DM, et al. ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res* 2014;42:D980-5.
- Xue Y, Chen Y, Ayub Q, Huang N, Ball EV, Mort M, et al. Deleterious- and Disease-Allele Prevalence in Healthy Individuals: Insights from Current Predictions, Mutation Databases, and Population-Scale Resequencing. *American Journal of Human Genetics* 2012;91:1022-32.
- Sim NL, Kumar P, Hu J, Henikoff S, Schneider G, Ng PC. SIFT web server: predicting effects of amino acid substitutions on proteins. *Nucleic Acids Res* 2012;40:W452-7.
- Choi Y and Chan AP. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics* 2015; 31:2745-7.
- Stone EA and Sidow A. Physicochemical constraint violation by missense substitutions mediates impairment of protein function and disease severity. *Genome Res* 2005;15:978-86.
- Tavtigian SV, Byrnes GB, Goldgar DE, Thomas A. Classification of rare missense substitutions, using risk surfaces, with genetic-and molecular-epidemiology applications. *Hum Mutat* 2008;29:1342-54.
- Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nat Methods* 2010;7:248-9.
- Li B, Krishnan VG, Mort ME, Xin F, Kamati KK, Cooper DN, et al. Automated inference of molecular mechanisms of disease from amino acid substitutions. *Bioinformatics* 2009;25:2744-50.
- De Baets G, Van Durme J, Reumers J, Maurer-Stroh S, Vanhee P, Dopazo J, et al. SNPeffect 4.0: on-line prediction of molecular and structural effects of protein-coding variants. *Nucleic Acids Res* 2012;40:D935-9.
- Valdar WS. Scoring residue conservation. *Proteins* 2002;48:227-41.
- Oliveira RA, Oshiro M, Hirata MH, Hirata RD, Ribeiro GS, Medeiros TM, et al. A novel point mutation in a class IV glucose-6-phosphate dehydrogenase variant (G6PD São Paulo) and polymorphic G6PD variants in São Paulo State, Brazil. *Genet Mol Biol* 2009;32:251-4.
- Hirono A, Fujii H, Hirono K, Kanno H, Miwa S. Molecular abnormality of a Japanese glucose-6-phosphate dehydrogenase variant (G6PD Tokyo) associated with hereditary non-spherocytic hemolytic anemia. *Hum Genet* 1992;88:347-8.
- Beutler E, Westwood B, Prchal J, Vaca G, Bartsocas CS, Baronciani L. New glucose-6-phosphate dehydrogenase mutations from various ethnic groups. *Blood* 1992;80:255-6.
- Beutler E, Westwood B, Melemed A, Dal Borgo P, Margolis D. Three new exon 10 glucose-6-phosphate dehydrogenase mutations. *Blood Cells Mol Dis* 1995;21:64-72.
- Hirono A, Kuhl W, Gelbart T, Forman L, Fairbanks VF, Beutler E. Identification of the binding domain for NADP+ of human glucose-6-phosphate dehydrogenase by sequence analysis of mutants. *Proc Natl Acad Sci U S A* 1989;86:10015-7.
- Zarza R, Pujades A, Rovira A, Saavedra R, Fernandez J, Aymerich M, et al. Two new mutations of the glucose-6-phosphate dehydrogenase (G6PD) gene associated with haemolytic anaemia: clinical, biochemical and molecular relationships. *Br J Haematol* 1997;98:578-82.
- Bulliamy T, Luzzatto L, Hirono A, Beutler E. Hematologically important mutations: glucose-6-phosphate dehydrogenase. *Blood Cells Mol Dis* 1997;23:302-13.
- Minucci A, Moradkhani K, Hwang MJ, Zuppi C, Giardina B, Capoluongo E. Glucose-6-phosphate dehydrogenase (G6PD) mutations database: review of the "old" and update of the new mutations. *Blood Cells Mol Dis* 2012;48:154-65.
- Kaczorowska-Hac B, Burzynska B, Plochocka D, Zak-Jasinska K, Rawa K, Adamkiewicz-Drozynska E. The first reported case of G6PD deficiency due to Seoul mutation in Poland. *Ann Hematol* 2014;93:879-80.
- Edison ES, Melinkeri SR, Chandy M. A novel missense mutation in glucose-6-phosphate dehydrogenase gene causing chronic nonspherocytic hemolytic anemia in an Indian family. *Ann Hematol* 2006;85:879-80.
- Filosa S, Cai W, Galanello R, Cao A, De Mattia D, Schettini F, et al. A novel single-base mutation in the glucose 6-phosphate dehydrogenase gene is associated with chronic non-spherocytic haemolytic anaemia. *Hum Genet* 1994;94:560-2.
- Mason PJ, Bautista JM, Gilsanz F. G6PD deficiency: the genotype-phenotype association. *Blood Rev* 2007;21:267-83.
- Weimer TA, Salzano FM, Westwood B, Beutler E. Molecular characterization of glucose-6-phosphate dehydrogenase variants from Brazil. *Hum Biol* 1993;41-7.

31. Sukumar S, Mukherjee MB, Colah RB, Mohanty D. Molecular characterization of G6PD Insuli—a novel 989 CGC → CAC (330 Arg → His) mutation in the Indian population. *Blood Cells Mol Dis* 2003;30:246-7.
32. Vaca G, Arámbula E, Monsalvo A, Medina C, Nuñez C, Sandoval L, et al. Glucose-6-phosphate dehydrogenase (G-6-PD) mutations in Mexico: four new G-6-PD variants. *Blood Cells Mol Dis* 2003;31:112-20.