Indonesian Journal of Biotechnology, June, 2013 Vol. 18, No. 1, pp.14-25

Phylogeny and Origin of Glucose-6-Phosphate Dehydrogenase (G6PD) Deficiency Mutations in Indonesia

Maria Omega[1] and Ross T. Barnard[1*]

[1] Biotechnology Program, School of Chemistry and Molecular Biosciences, The University of Queensland, Australia

Abstract

The aim of this study is to analyze the relationship between the types of G6PD mutations found in Indonesia and the relationships of mutations found in Indonesia to those found in other countries. We summarize the distribution of G6PDs in West Indonesia and East Indonesia. Moreover, we use bioinformatics methods to construct phylogenetic trees and compare the sequences containing the regions amplified by the commonly used PCR primer pairs. Previous work has shown that Mediterranean G6PD and Chinese CoimbraG6PD are distributed in West Indonesia, whilst G6PD mutations in East Indonesia are Jammu/ViangchanG6PD and Chinese Gaohe G6PD. G6PD Jammu/Viangchan was mostly distributed in Flores Island, East Indonesia along with G6PDGaohe. We constructed phylogenetic trees using the G6PD sequences from various regions in Indonesia and other countries. It appears from phylogenetic trees and percentages of identity that Flores Indonesian G6PD deficiency (Jammu/Viangchan G6PD, originating in India) is 92.5% identical to the G6PD deficiency of Chinese origin (GaoheG6PD). It was interesting to note that the genetic region containing the Javanese Indonesian G6PD deficiency (MediterraneanG6PD, first found in Italy) located in the western parts of Indonesia is closely related (99% identity) to the Chinese G6PD deficiency(Coimbra G6PD). We conclude that G6PD mutations in West Indonesia are closely related to G6PD mutations from China. G6PD mutations in East Indonesia are also closely related to G6PD mutations from India and China, but more distantly, and to different types to those in West Indonesia. A prediction of protein structure was carried out which allowed visualization of the locations of mutation on the three dimensional structure of G6PD.

Key words: G6PD, phylogeny, origin, genetic mutations

Introduction

Glucose-6-phosphate dehydrogenase (G6PD) deficiency is a metabolic enzyme involved in the pentose phosphate pathway, particularly vital in red blood cell metabolism. Frank (2005) found thatG6PD deficiency is the most common human enzyme defect, which is an X-linked recessivehereditary disease characterized by abnormally low levels of glucose-6-phosphate dehydrogenase. Moreover, G6PD Deficiency is now known to be slightly more common than previously thought in women though the symptoms are generally not as severe unless both X-chromosomes are affected. When it was discovered that G6PD Deficiency was a genetic disorder by Beutler *et al*. (1991) it wasthought to be a recessive disorder. Since women have two copies of the X-chromosome, they believed women would not be affected unless both chromosomes had the G6PD Deficiency mutation. The current view is that symptomatic patients are almost exclusively male, due to the X-linked pattern of inheritance, but female carriers can be clinically affected due to unfavorable lyonization, where random inactivation of an X-chromosome in certain cells creates a population of G6PD-deficient red blood cells

***Corresponding author :**
Ross T. Barnard
The University of Queensland, Brisbane QLD 4072, Australia phone: +61-07-33654612
email: rossbarnard@uq.edu.au

coexisting with normal red cells (Hamilton *et al*., 2004)

Mazza (2001) revealed that there are 440 different variants and the G6PD classification is divided by five types:

1. Chronic haemolytic anemia (<10% activity with severe deficiency).
2. Intermittent haemolysis (<10% activity with severe deficiency).
3. Haemolysis with stressors (10-60% activity with mild deficiency).
4. No clinical sequelae (non-deficient variant).
5. No clinical sequelae (increased enzyme activity).

Cappellini and Fiorelli (2008) showed that G6PD is the most common human enzyme defect, being present in more than 400 - 600 million people worldwide (8.96% of the world population). African, Middle Eastern/Mediterranean, European and South East Asian people are affected the most, along with those who are mixed with any of the above. Mehta *et al*. (2000) studied that a side effect of this disease is that it confers protection against malariain particular the form of malaria caused by *Plasmodium falciparum*, the most deadly form of malaria. A similar relationship exists between malaria and sickle-cell disease. This is because the cells infected with the *Plasmodium* parasite are cleared more rapidly by the spleen (Monga *et al*.,2003). This phenomenon might give G6PDH deficiency carriers an evolutionary advantage by increasing their fitness in malarial endemic environments. Specifically, Indonesia carries one of the highest burdens of malaria in the world, and Primaquine (PQ) as the only available drug for radical cure of *Plasmodium vivax* and prevention of*P. falciparum* transmission. The PQ can trigger acute haemolysis in G6PD deficient individuals and cause problems in malaria treatment. Furthermore, Satyagraha (2012) reported thatin Indonesia, G6PD rates are relatively high in the many different and geographically isolated ethnic groups. The background of the study is that G6PD 1388 G>A (known as Kaiping), which is a variant suggested to be of Chinese origin, is found in Southeast Asia. G6PD Kaiping is widely distributed among native Flores and Palue Islanders for unknown reasons, while G6PD Canton (1376 G>T) is very rare or entirely absent in these locations (Tantular *et al*., 2010). Therefore, it is interesting to investigate whether these variants or other found in Indonesia arose by local mutations or were introduced by Chinese or other migrants. To elucidate the answers, bioinformatics analysis is needed, such as a haplotype study of Xq28 of Chinese, Indian and Indonesian people. In the present study, we used bioinformatics to investigate the relationships between the mutations found in Indonesia and those found in other countries.

Materials and Methods

Sequence data was compared to the GenBank DNA Database using BLASTn searches to determine the alignment (% identity) between primers and sequences containing G6PD mutations using the National Centre for Biotechnology Information (NCBI) BLAST network server available from http://www.ncbi.nlm.nih.gov/

Construction of phylogenetic trees.

The materials are the computers and the software such as ClustalX v 2.1 to construct phylogenetic trees. The phylogenetic trees were developed to test the relationship between regions of sequence containing the squences bounded by PCR primers which were aligned against a selection of genbank database sequences using ClustalX v 2.1 available from http://www.clustal.org/.We used G6PD mutations that cause G6PD deficiency. They are located on the long arm of X chromosome, on band Xq28 (McKusick and Hamosh, 2000). The G6PD gene spans 18.5 kilobases (Warrel *et al*., 2005). Sequences based on databases were uploaded as FASTA files and a preliminary complete alignment was done. Regions containing sequences defined by primers

were aligned to draw bootstrapped NJ trees (using 1000 bootstrap steps) and the trees were visualized using Treeview.

Results

Sets of primers to identify the G6PD deficiency and mutations

There are 14 sets of primers that have been used previously to identify the G6PD deficiency in Indonesian, Chinese and Indian people (Table 1, 2 and 3).

There are 13 G6PD sequences from NCBI used in this report (Table 4).

Mutations of G6PD deficiency

G6PD protein was submitted to http://zhanglab.ccmb.med.umich.edu/I-TASSER/output/S110585

Phylogenetic trees

Phylogenetic tree (Figure 3) was generated based on a block of 910 bp that includes mutations at 1376 and at 392 in adjacent regions 3 and 4 of G6DP on chromosome Xq28 (Table 8).

Table 1. Different Sets of G6PD primers used in Indonesian people

G6PD mutation	Sense/ antisense strand	G6PD primers (nucleotides)	Restriction Enzymes	Design Locations	References
nt1376 (G→T)	Sense	5′-ACGTGAAGCTCCTGACGC-3′	XhoI	Indonesia China	Matsuoka *et al.* (2003) Kawamoto *et al.* (2006) Soemantri *et al.* (1995)
	Antisense	5′-TGAAAATACGCCAGGCCTCG -3′			
nt1388 (G→A)	Sense	5′-ACGTGAAGCTCCTGACGC-3′	NdeI	Indonesia China	Matsuoka *et al*., (2003) Kawamoto *et al*., (2006) Soemantri *et al*., (1995) Tantular *et al*., (2010)
	Antisense	5′-GTGCAGCAGTGGGGTGAA CATA-3′			
nt95 (A→G)	Sense	5′-CGTTCACAAGGAGTGATTTG-3′	MluI	Indonesia China	Matsuoka *et al*. (2007), Tantular *et al.* (2010)
	Antisense	5′-CGATGCACCCATGATGAT GAATACG-3′			
nt1024(C→T)	Sense	5′-GTCAAGGTGTTGAAATGCA TC -3′	MboII	Indonesia China	Matsuoka *et al.* (2003), Kawamoto *et al.* (2006)
	Antisense	5′-CATCCCACCTCTCATTCTT CC-3′			
nt487(C→T)	Sense	5′-GCGTCTGAATGATGCTGCT GTGAT-3′	AluI	Indonesia India	Matsuoka *et al.* (2003), Kawamoto *et al.* (2006)
	Antisense	5′-AGCCGGTCAGTGCTCTGC ATGTCC-3′			

Table 2. Different Sets of G6PD primers used in Chinese people (Matsuoka et al., 2007)

G6PD mutation	Sense/antisense	G6PD Primers	Nucleotide size in basepairs (bp)
nt175(C→T)	Sense	5′-ACATGTGGCCCCTGCACCACA-3′	242 bp
	Antisense	5′-GTGACTGCTCTGCCACCCTG-3′	
nt163(C→T)	Sense	5′-TTGGGGTCCCCATGCCCTTG-3	231 bp
	Antisense	5′-TGCCTCGTCACAGATGGGCC-3′	
nt1376(G→T)	Sense	5′-CTGAGAGAGCTGGTGCT-3′	342 bp
	Antisense	5′-CACCATGTGGAGTCCCCCGG-3′	
nt1311(C→T)	Sense	5′-ACTTCCACATGGTGGCAGGCAG-3′	397 bp
	Antisense	5′-ATGAGGTAGCTCCACCCTCA-3′	
nt1388(G→A)	Sense	5′-TGAGGGTGGAGCTACCTCAT-3′	164 bp
	Antisense	5′-CGGGGGTGGAGGTGGGTGCCCA-3′	

Table 3. Different Sets of G6PD primers used in Indian people (Vulliamy *et al*., 1991)

G6PD mutation	Sense/antisense	G6PD Primers	Restriction enzyme
nt175(C→T)	Sense	5′-TGGACCCCTACACAGCCAAGTAC-3′	ScaI and SacI
	Antisense	5′-GGCATGCTCCTGGGGACTGCT-3′	
nt163(C→T)	Sense	5′-GGAGCTAAGGCGAGCTCTGGC-3′	BspHI
	Antisense	5′-TGCCTTGCTGGGCCTCGAAGG-3′	
nt563 (C→T)	Sense	5′-ACTCCCGAAGAGGGGTTCAAGG-3′	MboII
	Antisense	5′-CCAGCCTCCCAGGAGAGAGGAAG-3′	
nt1311(C→T)	Sense	5′-TGTTCTTCAACCCCGAGGAGT-3′	BclI
	Antisense	5′-AAGACGTCCAGGATGAGGTGATC-3′	

Table 4. G6PD sequences, their origins and distributions

No	G6PD and its sequences in NCBI	Distributions	G6PD mutation and names
1.	>gi|108773792|ref|NM_001042351.1| Homo sapiens glucose-6-phosphate dehydrogenase (G6PD), transcript variant 2, exon 11, mRNA **Origins:** Kobe, Japan; Telti, Italy (Mason *et al*., 1995)	Israel, Jordan, Cairo, Mediterranean, Uzbekistan, India, China	nt1318 (C>T) Telti
2.	>gi|108773794|ref|NM_000402.3|Homo sapiens glucose-6-phosphate dehydrogenase (G6PD), transcript variant 1, exon 11, mRNA **Origins:** Britain, UK; Iran (Mortazavi *et al*., 1997)	Israel,Jordan, Cairo, Mediterranean, Uzbekistan, India, China	nt1311 (C>T) Iranian/European
3.	>gi|213385264|ref|NG_009015.1| Homo sapiens glucose-6-phosphate dehydrogenase (G6PD), RefSeqGene (LRG_148) on chromosome X, exon 9 **Origins:**Gaozhou, China (Chao, *et al*., 1991).	Germany	nt95 (A>G) Gaohe
4.	>gi|122719247|gb|EF190463.1| Homo sapiens glucose-6-phosphate dehydrogenase (G6PD) gene, exon 6 and partial cds **Origins:** Southern China; Mahidol, Thailand (Vulliamy *et al*., 1989).	China	nt487(G>A) Mahidol
5.	>gi|215769502|dbj|AB376963.1| Homo sapiens G6PD gene for glucose-6-phosphate dehydrogenase variant, exon 8, partial cds **Origins:** Jammu, India; Gifu, Japan (Beutler *et al*., 1991).	Papua New Guinea, East Indonesia, Japan	nt871 (G>A) Jammu/Viangchan
6.	>gi|112818958|gb|DQ839546.1| Homo sapiens glucose-6-phosphate dehydrogenase variant (G6PD) gene, exon 5 and partial cds **Origins:**Sassari/Cagliari, Italy; Panama, America (Beutler *et al*., 1990)	Japan, India, West Indonesia	nt563 (C>T) Mediterranean
7.	>gi|111052656|gb|DQ832765.1| Homo sapiens glucose-6-phosphate dehydrogenase (G6PD) gene, exons 3 and partial cds **Origins:** China (Chiu *et al*., 1993)	China	nt392 (G>T) Chinese-4
8.	>gi|111052652|gb|DQ832763.1| Homo sapiens glucose-6-phosphate dehydrogenase (G6PD) gene, exons 12 and partial cds **Origins:** Taiwan (Tang *et al*., 1992)	China	nt493 (A>G) Chinese-3
9.	>gi|111052648|gb|DQ832761.1| Homo sapiens glucose-6-phosphate dehydrogenase (G6PD) gene, exon 7 and partial cds **Origins:**Guangzhou, Southern China (Ren *et al*., 1999)	China	nt175 (C>T) IVS

No	G6PD and its sequences in NCBI	Distributions	G6PD mutation and names
10.	>gi|111052658|gb|DQ832766.1| Homo sapiens glucose-6-phosphate dehydrogenase (G6PD) gene, exons 4 and partial cds **Origins:**Guangzhou, China; Hakka, Taiwan (Stevens *et al*., 1990)	China	nt1376 (G>T) Canton
11.	>gi|111052654|gb|DQ832764.1| Homo sapiens glucose-6-phosphate dehydrogenase (G6PD) gene, exon 13, partial cds **Origins:** Guangdong, China; Anand/Dhon India (Chiu *et al*., 1991)	China	nt1388 (G>A) Kaiping
12.	>gi|111052650|gb|DQ832762.1| Homo sapiens glucose-6-phosphate dehydrogenase (G6PD) gene, exon 5 and partial cds **Origins:** Coimbra, Portugal; Shunde, Guangdong, China (Du *et al*., 1992)	China	nt592 (C>T) Coimbra
13.	>gi|31538|emb|X53815.1| Human G6PD gene for glucose-6-phosphate dehydrogenase, exon 5, partial. **Origins:**Shinshu, Japan; Agrigento, Italy (Hirono *et al*., 1994a,b)	Italy	nt527 (A>G) Shinshu

Table 5. Primers with high identity used to detect G6PDd of Indonesia, China and India

Origins	References	% Identity	G6PD Primers
India	Vulliamy *et al*., (1991)	71.3%	nt1376-F (5′-ACGTGAAGCTCCTGACGC-3′)
		82.2%	nt1388-R (5′-GTGCAGCAGTGGGGTGAACATA-3′)
Indonesia	Matsuoka *et al*., (2003)	94%	nt1388-F (5′-TGAGGGTGGAGCTACCTCAT-3′)
	Kawamoto *et al*., (2006)	96.9%	nt1376-R (5′-CACCATGTGGAGTCCCCCGG-3′)
	Soemantri *et al*., (1995)		
	Tantular *et al*., (2010)		
China	Matsuoka *et al*., (2007)	70.9%	nt163-F (5′-GGAGCTAAGGCGAGCTCTGGC-3′)
		70.7%	nt1311-R (5′-AAGACGTCCAGGATGAGGTGATC-3′)
Indonesia, India, China	Matsuoka *et al*., (2007)	**58.7%**	nt1376-F (5′-CTGAGAGAGCTGGTGCT-3′)
		63.9%	nt1311-R (5′-ATGAGGTAGCTCCACCCTCA-3′)

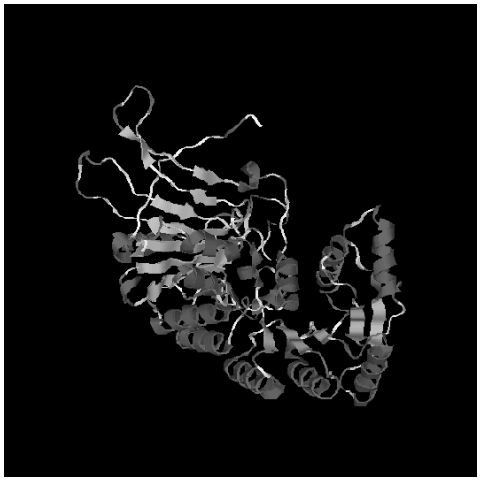
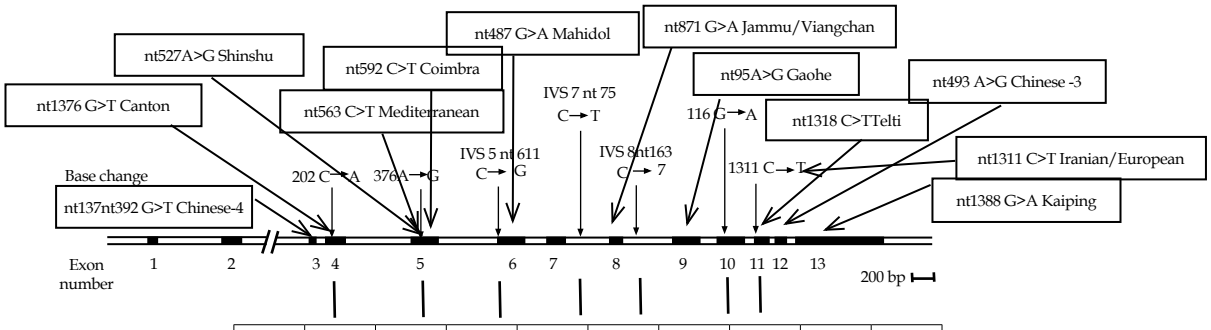


Figure 1. Predicted G6PD protein structure (3.6±2.5Å RMSD from i-TASSER Michigan University)

Table 6. Haplotypes of G6PD genes in different polymorphic sites (Vulliamy *et al*., 1991)

Characteristics	G6PD genes	G6PD genes	G6PD gene
Base change	nt175 C>T	nt163 C>T	nt1311 C>T
Exon number	7	8	11
Restriction sites	ScaI	BspHI	BclI
Distributions	Chinese and Indian	Chinese and Indian	Iranian/European, Javanese-Indonesian and Indian
Haplotypes	A and A-	B and A	B



Restriction site / Haplotype number

	NlaIII	Fold	PvulI	ScaI	BspHI	PstI	BcLI	%	G6PD
I	-	-	-	-	+	+	-	20	B
VII	-	-	-	-	+	+	+	24	B
II	-	-	-	-	+	-	-	11	B
III	-	+	-	-	+	+	-	2	A
IV	-	+	-	-	-	+	-	13	A
V	-	+	+	+	-	+	-	4	A
VI	+	+	+	+	-	+	-	26	A

Figure 2. Haplotypes of G6PD and locations of polymorphic sites (Vulliamy *et al*., 1991)

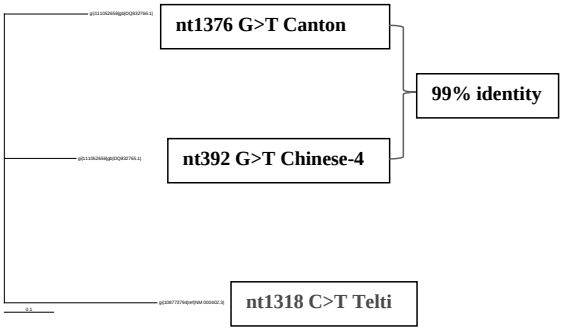


Figure 3. nt392 G>T Chinese-4 and nt1376 G>T Canton (99% high identity) bounded by Chinese primers of nt1376-F (5′-CTGAGAGAGCTGGTGCT-3′) and nt1311-R (5′-ATGAGGTAGCTCCACCCTCA-3′) with the outgroup of nt1318 C>T Telti. The PCR product is 90 bp and the block of aligned sequence is 910 bp.

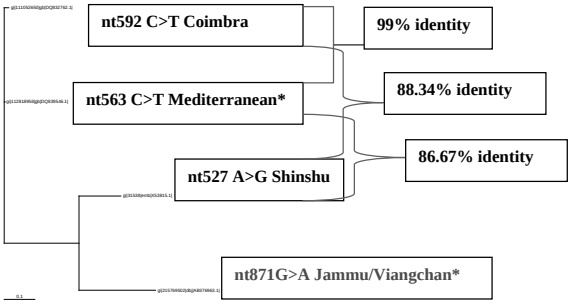


Figure 4. nt563 C>T Mediterranean and nt592 C>T Coimbra (99% high identity), nt527 A>G Shinshu and nt563 C>T Mediterranean (86.67% moderate identity), nt527 A>G Shinshu and nt592 C>T Coimbra (88.34% moderate identity), bounded by Chinese primers of nt1376-F (5′-CTGAGAGAGCTGGTGCT-3′) and nt1311-R (5′-ATGAGGTAGCTCCACCCTCA-3′) with the outgroup of nt871 G>A Jammu/Viangchan. The PCR product is 60 bp and the block of aligned sequence is 580 bp.Asterisk (*) is G6PD mutation located in Indonesia

Table 7. G6PD mutated genes and regions inchromosome (Xq28)(http://www.ncbi.nlm.nih.gov/)

Regions	G6PD mutated genes	NCBI	% Identity
Region 3 Region 4	nt392 G>T Chinese-4 nt1376 G>T Canton* (Figure 3)	>gi|111052656|gb|DQ832765.1| >gi|111052658|gb|DQ832766.1	99% between a 910 bp region enclosing Chinese-4 and Canton = High identity
Region 5	nt527 A>G Shinshu **nt563 C>T Mediterranean*** nt592 C>T Coimbra (Figure 4)	>gi|31538|emb|X53815.1| >gi|112818958|gb|DQ839546.1| >gi|111052650|gb|DQ832762.1|	• 99% between a 580 bp region enclosing Coimbra andMediterranean = High identity • 88.34% between a 580 bp region enclosingShinshu and Coimbra = Moderate identity • 86.67% between a 580 bp region enclosingShinshu and Mediterranean = Moderate identity
Region 6 Region 7	nt487 G>A Mahidol nt175 C>T IVS (Figure 5)	>gi|122719247|gb|EF190463.1| >gi|111052648|gb|DQ832761.1|	74.07% between a 520 bp region enclosing Mahidol and IVS = Low identity
Region 8 Region 9	**nt871 G>A Jammu/ Viangchan*** nt95 A>G Gaohe (Figure 6)	>gi|215769502|dbj|AB376963.1| >gi|213385264|ref|NG_009015.1|	92.5% similarity between a 720 bp region enclosing Jammu/Viangchan and Gaohe= Moderate identity
Region 11	nt1318 C>T Telti nt1311 C>T Iranian/ European (Figure 7)	>gi|108773792|ref|NM_0010423 51.1| >gi|108773794|ref|NM_000402.3|	99% similarity between a 550 bp region enclosing Telti and Iranian/European = High identity
Region 12 Region 13	nt493 A>G Chinese -3 nt1388 G>A Kaiping* (Figure 8)	>gi|111052652|gb|DQ832763.1| >gi|111052654|gb|DQ832764.1|	74.12% betweena 830 bp region enclosing Chinese-3 and Kaiping= Low identity

Note: * (asterisk) indicates the mutated G6PD genes located in Indonesia

Phylogenetic tree (Figure 4) was generated based on a block of 580 bp that includes mutations at 527, at 563 and at 592 in adjacent region 5 of G6DP on chromosome Xq28 (Table 8).

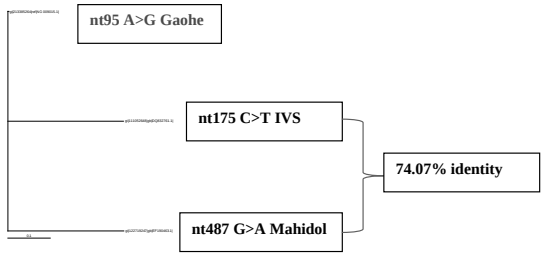


Figure 5. nt487 G>A Mahidol and nt175 C>T IVS (74.07% low identity), bounded by Chinese primers of nt1376-F (5′-CTGAGAGAGCTGGTGCT-3′) and nt1311-R (5′-ATGAGGTAGCTCCACCCTCA-3′) with the outgroup of nt95 A>G Gaohe. The PCR product is 120 bp and the block of aligned sequence is 520 bp.

Phylogenetic tree (Figure 5) was generated based on a block of 520 bp that includes mutations at 175 and at 487 in adjacent regions 6 and 7 of G6DP on chromosome Xq28 (Table 8).

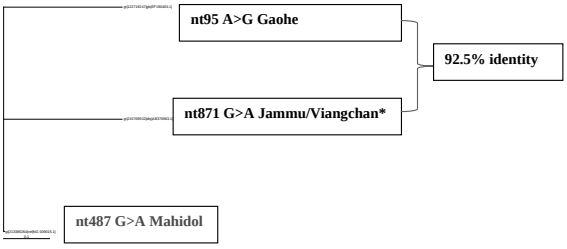


Figure 6. nt871 G>A Jammu/Viangchan and nt95 A>G Gaohe (92.5% moderate identity), bounded by Chinese primers of nt1376-F (5′-CTGAGAGAGCTGGTGCT-3′) and nt1311-R (5′-ATGAGGTAGCTCCACCCTCA-3′) with the outgroup of nt487 G>A Mahidol. The PCR product is 80 bp and the block of aligned sequence is 720 bp. Asterisk (*) is G6PD mutation located in Indonesia.

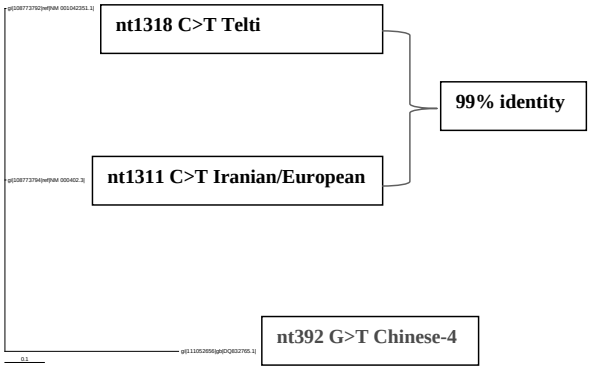


Figure 7. nt1318 C>T Telti and nt1311 C>T Iranian/ European (99% high identity), bounded by Chinese primers of nt1376-F (5′-CTGAGAGAGCTGGTGCT-3′) and nt1311-R (5′-ATGAGGTAGCTCCACCCTCA-3′) with the outgroup of nt392 G>T Chinese-4. The PCR product is 90bp and the block of aligned sequence is 550 bp.

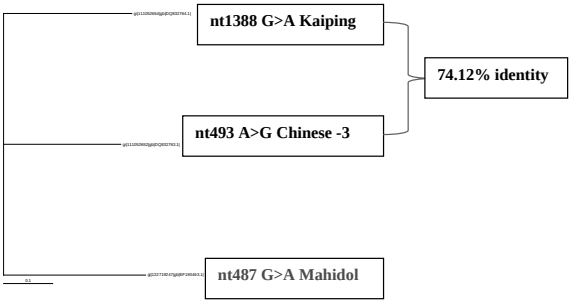


Figure 8. nt493 A>G Chinese -3 and nt1388 G>A Kaiping (74.12% low identity), bounded by Chinese primers of nt1376-F (5′-CTGAGAGAGCTGGTGCT-3′) and nt1311-R (5′-ATGAGGTAGCTCCACCCTCA-3′) with the outgroup of nt487 G>A Mahidol. The PCR product is 190 bp and the block of aligned sequence is 830bp.Asterisk (*) is G6PD mutation originally located in Indonesia.

Phylogenetic tree (Figure 6) was generated based on a block of 720 bp that includes mutations at 95 and at 871 in adjacent regions 8 and 9 of G6DP on chromosome Xq28 (Table 7).

Phylogenetic tree (Figure 7) was generated based on a block of 550 bp that includes mutations at 1311 and at 1318 in adjacent region 11 of G6DP on chromosome Xq28 (Table 7).

Phylogenetic tree (Figure 8) was generated based on a block of 830 bp that includes mutations at 493 and at 1388 in adjacent regions 12 and 13 of G6DP on chromosome Xq28 (Table 7).

Discussion

The predicted protein structure of G6PD mutation showed a monomer protein with a substrate binding cavity between two domains (Figure 1). The locations of amino acid mutations are at the binding sites, presumed to have functions in protection against malaria (Mason, *et al*., 2007). The polymorphic mutations decrease stability of the enzyme in red cells, by disturbing protein folding, affecting amino acid residues at dimer interface or interacting with a structural NADP molecule that stabilizes the enzyme (Mason, *et al*., 2010).As discussed by Turner *et al*. (2001), a number of different G6PD variants have reached polymorphic frequencies in different parts of the world due to the relative protection they confer against malaria infection. Mason *et al*. (2007) showed that very rarely *de novo* mutations can arise causing the more severe condition of chronic nonspherocytic hemolytic anemia. Altogether 169 different mutations were described (Mason *et al*., 2010). Iwai *et al*. (2001) showed that G6PD mutations in Chinese are heterogeneous and at least 10 mutations have been reported including G6PD Canton and G6PD Kaiping. Kaiping mutations specifically showed minimal mobility shift. Interestingly, Iwai *et al*. (2001) Soemantri *et al*. (1995) also found that G6PD Kaiping was not found in orang asli (original people) from Indonesia, India, Malaysia and Myanmar, but was found in Chinese, Singaporean Chinese and Taiwan Chinese. G6PD Canton was found in Javanese Indonesian, Chinese, Singaporean Chinese and Taiwan Chinese. However, further study by Matsuoka *et al*. (2003) revealed that G6PD Kaiping was indeed found in Maumere and Telibura at Flores Island, Indonesia while G6PD Canton was found in Central Java, Indonesia. It is important to note that the history of Flores Island, in East Indonesia, is consistent with a diverse origin of mutations.

It belongs to the Sunda archipelago where people come from Eurasian countries, African countries, Philippine Islands, or Pacific islands. Therefore, historically, Flores Islands might have received many tribes from different origins. The high frequency distribution of G6PD Kaiping, a common variant in Chinese population, in the Sikka, Flores Islands is a rather unexpected finding (Kawamoto *et al*., 2006). In contrast, Jiang *et al*. (2006) characterized that G6PD Canton was never seen in the Sikka, Flores Islands nor in the Ende, Timor Islands, although many studies revealed that G6PD Canton in China, Taiwan, Thailand, Malaysia, and Singapore is found at higher than or almost equal in frequency to G6PD Kaiping in those countries. Bioinformatics is needed to discover the genetic relationships between mutations that are common in Indonesia, and those found in China, India and the Middle East. For instance, mutations at nt175, nt163 and nt1311 are summarized in Table 6. Haplotype number V and VI are thought to be from Chinese origins whereas haplotype number VII are from Javanese Indonesians and Indians, as studied by Vulliamy *et al*. (1991). Additionally, it was presumed by Kawamoto *et al*. (2006) that Chinese traders or immigrants have shared genes with Christian Sikka women and so introduced G6PD Kaiping mutated gene into the Sikka population in Flores Islands, Indonesia. However, Indonesian G6PD mutated genes also consist of G6PD Jammu/Viangchan in East Indonesia (including Flores Islands). It is clear from Soemantri *et al*. (1995) that G6PD deficiency is found in Java Island, Indonesia and is represented by the G6PD mutations of Mediterranean and G6PD mutations of Canton, consistent with a Mediterranean, Indian and Chinese ("Mongoloid" is the term used by Soemantri *et al*. (1995) contribution to the gene pool of the Javanese. G6PD Mediterranean in West Indonesia (including Java Island) is most closely related to Chinese G6PD mutated genes (G6PD Coimbra) with 99% identity (see Figure 4). From Table 6 and Figure 2, it can be seen that Flores Indonesian G6PD deficiency (Jammu/Viangchan) is moderately related (92.5% identity) to Chinese G6PD mutation (Gaohe) (see Table 7 and Figure 6). The phylogenetic tree of regions containing mutation bounded by Chinese PCR primers is shown in Figure 3, 4, 5, 6, 7 and 8. The block containing the G6PD mutations found in Indonesia(nt563 C>T Mediterranean) is highly related to Chinese G6PD (nt592 C>T Coimbra) with 99%identity between the regions containing these gene mutations(see Figure 4). Indonesian G6PD (Mediterranean) and Chinese G6PD (Coimbra) are located in the same exons (at exon 5) as well as possessing high identity in the sequence adjacent to the mutations. The G6PD mutation found in Indonesia (nt563 C>T Mediterranean) is moderately related to Chinese G6PD (which is also found in Italy, nt527 A>G Shinshu) with 86.67%identity (Figure 4). Both are located in exon 5. Chinese G6PD mutation (Shinshu)is moderately related to Chinese (Coimbra) G6PD mutation with 88.34% identity (see Figure 4). A phylogenetic tree of regions containing G6PD segments bounded by Chinese PCR primers showed 92.5% (moderate) identity (Figure 6) between G6PD found in Indonesia (nt871 G>A Jammu/Viangchan) and Chinese G6PD(nt95 A>G Gaohe). It can be concluded from the phylogenetic trees that G6PD Mediterranean in Indonesia is most closely related to G6PD Coimbra, while G6PD Jammu/Viangchan in Indonesia has less similarity with G6PDGaohe. The block of DNA bounding G6PD Mahidol is more distantly related to G6PD IVS with 74.07% (low identity)(see Figure 5), whereas G6PD Iranian/European has high identity (99%) with G6PD Telti, found in India (see Figure 7). It is important to note that G6PD Canton, found in Java Island, West Indonesia, is closely related to G6PD Chinese-4 with 99% high identity (see Figure 3). Meanwhile G6PD Kaiping, mostly distributed in Flores Island, East Indonesia is more distantly related to G6PD Chinese-3 with 74.12% identity (see

Figure 8). G6PD Kaiping is found not only in Flores Island but also in Java Island. G6PD Jammu/Viangchan is only found in Flores Island. In addition, there were other phylogenetic trees constructed for reviewing the best primers with high percentages of identity for amplifying samples from various locations. The primers with high percentages of identity for use with Indonesian samples are the Chinese primers (i.e. a pair of primers of nt1388-F and nt1376-R), and other Chinese primers of nt1376-F and nt1311-R (for Indonesian, Indian and Chinese samples). The Chinese primers of nt1376-F and nt1311-R for Indonesian, Indian and Chinese generate products that fall within the blocks of DNA used to generate phylogenetic trees for regions containing the mutations of the G6PD genes. The primers can be used to detect G6PD mutated gene samples, as summarized in Table 5. Based on our phylogenetic trees and the percentages of identity between blocks of DNA sequence enclosing the mutations (see Table 7), we conclude that the Flores Indonesian G6PD deficiency (Jammu/ Viangchan G6PD), which is located in the eastern parts of Indonesia is moderately (92.5%) related to the G6PD deficiency (GaoheG6PD) of Chinese origin, consistent with a distant Chinese origin for this mutation, but with significant evolutionary divergence between these regions, which is likely to have occurred prior to introduction of the Jammu/ Viangchan mutation into Indonesia. It was interesting to note that in the region containing the PCR product bounded by PCR primer pair of nt1376 and nt1311, theJavanese Indonesian G6PD deficiency (Mediterranean G6PD)is highly related (99%) to the Chinese G6PD deficiency (Coimbra G6PD), but only moderately related to Shinshu G6PD deficiency. It has been suggested that the originally reported G6PD mutations in Java Island (Jakarta, Solo and Surabaya) and Bali Island (Coimbra G6PD, Shinshu G6PDand Canton G6PD); as well as G6PD mutation is Flores Island and east Indonesian islands (Jammu/Viangchan G6PD, Gaohe G6PD and Kaiping G6PD) were introduced by Chinese migrants since the year 1293, whilst G6PD mutations in Sumatra Island (Medan) and Java Island (Jakarta and Bandung) (Mediterranean G6PD) were introduced byIndian traders since the first century (Brandes, 1913; Tantular, *et al.*, 2010). However, our results support the suggestion that a mutation in Java and west Indonesia (Mediterranean G6PD) is very closely related to Chinese G6PD deficiency (Coimbra G6PD), whilst the mutation originating in India (Jammu/Viangchan G6PD) and now distributed in east Indonesia, is moredistantly related to Gaohe G6PD deficiency of Chinese origin.

Acknowledgments

This work was funded by Australian Leadership Award from Ausaid. Thank you to Ari Satyagraha for help during this work.

References

Beutler, E., and Kuhl, W. 1990. The NT 1311 polymorphism of G6PD: G6PD Mediterranean mutation may have originated independently in Europe and Asia. *Am. J. Hum. Genet.*,**47**, 1008-1012.

Beutler, E., Prchal, J.T., Westwood, B., and Kuhl, W. 1991. Definition of the mutations of G6PD Wayne, G6PD Viangchan, G6PD Jammu and G6PD "LeJeune". *Acta Haematol.*, **86**, 179-182.

Brandes, J.L.A. (1913). Oud-Javaanscheoorkonden: nagelatentransscripties van wijlen J.L.A. Brandes ;uitgegeven door N.J. Krom. Batavia: Albrecht.

Cappellini, M.D., and Fiorelli, G., 2008. Glucose-6-phosphate dehydrogenase deficiency. *Lancet*, **371**(9606), 64–74.

Chao, L., Du, C.S., and Louie, E. 1991. A to G substitution identified in exon 2 of the G6PD gene among G6PD deficient Chinese. *Nucleic Acids Res.*, **19**, 6056.

Chiu, D.T.Y., Zuo, L., and Chao, L. 1993. Molecular characterization of glucose-

6-phosphate dehydrogenase (G6PD) deficiency in patients of Chinese descent and identification of new base substitutions in the human G6PD gene. *Blood*,**81**, 2150-2154.

Chiu, D.T.Y., Zuo, L., and Chen, E. 1991. Two commonly occurring nucleotide base substitutions in Chinese G6PD variants. *Biochem.Biophys. Res.Commun*., **180**,988-993.

Clustal-X., 2012. Clustal-X v 2.1.Available from http://www.clustal.org/ on 21 July 2012.

Du, C.S., Chao, L.T., Louie, E., Liu, T.Z., and Chiu, D.T.Y. 1992.Molecular characterization of G6PD deficiency in patients of Chinese descent and identification of a new base substitution in the human G6PD gene.*Blood,* **80** (Suppl 1), 284a.

Frank, J. E., 2005. Diagnosis and management of G6PD deficiency. *Am. Fam. Physician*., **72**(7), 1277-82.

Hamilton, J. W., Jones, F. G., and McMullin, M. F., 2004. G6PD Guadalajara: a case of chronic non-spherocytichaemolyticanaemia responding to splenectomy and the role of splenectomy in this disorder. *Hematology*,**9**(4), 307–9.

Hirono, A., Miwa, S., Fujii, H., Ishida, F., Yamada, K., Kubota, K. 1994a.Molecular study of eight Japanese cases of G6PD by nonradio isotopic single-strand conformation polymorphism analysis. *Blood*,**83**,3363-3368.

Hirono, A., Nakayama, S., Fujii, H., and Miwa, S. 1994b. Molecular abnormality of a unique Japanese G6PD variant (G6PD Kobe) with a greatly increased affinity for galactose-6-phosphate. *Am. J. Hematol*.,**45**, 185-186.

i-TASSER., 2012. i-TASSER results for job id S110585. Available from http://zhanglab.ccmb.med.umich.edu/I-TASSER/output/S110585 on 21 August 2012.

Iwai, K., Hirono, A., Matsuoka, H., Kawamoto, F., Horie, T., Lin, K., Tantular, I. S., Dachlan, Y. P., Notopuro, H., Hidayah, N. I., Salim, A. M. A., Fujii, H., Miwa, S., and Ishii, A., 2001. Distribution of glucose-6-phosphate dehydrogenase mutations in Southeast Asia. *Hum. Genet*., **108**, 445-449.

Jiang, W., Yu, G., Liu, P., Geng, Q., Chen, L., Lin Q., Ren, X., Ye, W., He, Y., Guo Y., Duan, S., Wen, J., Li, H., Qi, Y., Jiang, C., Zheng, Y., Liu, C., Si, E., Zhang, Q., Tian, Q., and Du, C., 2006. Structure and function of glucose-6-phosphate dehydrogenase-deficient variants in Chinese population. *Hum. Genet*., **119**, 463-478.

Kawamoto, F., Matsuoka, H., Kanbe, T., Tantular, I. S., Pusarawati, S., Kerong, H. I., Damianus, W., Mere, D., and Dachlan, Y. P., 2006. Further investigation of glucose-6-phosphate dehydrogenase variants in Flores Island, eastern Indonesia. *J. Hum. Genet*., **51**, 952-957.

Mason, P. J., Bautista, J. M., and Gilsanz, F., 2007. G6PD deficiency: the genotype-phenotype association. *Blood Rev*., **21**(5), 267-283.

Mason, P. J., Bautista, J. M., and Gilsanz, F., 2010. Corrigendum to G6PD deficiency: the genotype-phenotype association. *Blood Rev*., **24**(1), 49.

Mason, P.J., Sonati, M.F., and MacDonald, D. 1995. New glucose-6-phosphate dehydrogenase mutations associated with chronic anemia. *Blood*. **85**, 1377-1380.

Matsuoka, H., Arai, M., Yoshida, S., Tantular, I. S., Pusarawati, S., Kerong, H., and Kawamoto, F., 2003. Five different glucose-6-phosphate dehydrogenase (G6PD) variants found among 11 G6PD-deficient persons in Flores Island, Indonesia. *J. Hum. Genet*., **48**, 541-544.

Matsuoka, H., Thuan, D. T. V., Thien, H. V., Kanbe, T., Jalloh, A., Hirai, M., Arai, M., Dung, N., T., and Kawamoto, F., 2007. Seven different glucose-6-phosphate dehydrogenase variants including a new variant distributed in Lam Dong province in Southern Vietnam. *Acta. Med. Okayama*., **61**(4), 213-219.

Mazza, J. (2001). Manual of clinical hematology. New York: Lippincott Williams and Wilkins.

McKusick, V, A., and Hamosh, A., 2010. Anemia, nonspherocytic hemolytic, due to G6PD deficiency.Available from http://omim.org/entry/305900

Mehta, A., Mason, P. J., and Vulliamy, T. J., 2000. Glucose-6-phosphate dehydrogenase deficiency.*Baillieres Best Pract. Res. Clin. Haematol.*,**13**(1), 21–38.

Monga, A., Makkar, R.P., Arora, A., Mukhopadhyay, S., and Gupta, A.K., 2003.Case report: Acute hepatitis E infection with coexistent glucose-6-phosphate dehydrogenase deficiency. *Can. J. Infect. Dis.*,**14**(4), 230–1.

Mortazavi, Y.,Chopra, R., Gordon-Smith, E.C., and Rutherford, T.R. 1997. Frequency of the G6PD nt 1311 C/T polymorphism in English and Iranian populations: relevance to studies of X chromosome inactivation. *J Med Genet.*, **34**(12), 1028-9.

NCBI., 2012. G6PD and Homo sapiens. Available from http://www.ncbi.nlm.nih.gov/nuccore

Ren, X.,Du, C., and Lin, Q. 1999. Studies on a G6PD polymorphic site, C1311T. *ZhonghuaXue Ye XueZaZhi*, **20**(4), 197-9.

Satyagraha, A. W. (2012). G6PD Testing in Indonesia. Incheon: Asia Pacific Malaria Elimination Network (APMEN).

Soemantri, A. G., Saha, S., Saha, N., and Tay, J. S. H., 1995. Molecular variants of red cell glucose-6-phosphate dehydrogenase deficiency in Central Java, Indonesia. *Hum. Hered.*, **45**, 346-350.

Stevens, D.J., Wanachiwanawin, W., Mason, P.J., Vulliamy, T.J., and Luzzatto, L. 1990. G6PD Canton a common deficient variant in South East Asia caused by a 459 Arg->Leu mutation. Nucleic Acids Res, 18, 7190.

Tang, T.K., Huang, C.S., Huang, M.J., Tam, K.B., Yeh, C.H., and Tang, C.J.C. 1992. Diverse point mutations result in glucose 6-phosphate dehydrogenase (G6PD) polymorphism in Taiwan. *Blood*, **79**, 2135-2140.

Tantular, I. S., Matsuoka, H., Kasahara, Y., Pusarawati, S., Kanbe, T., Tuda, J. S. B., Kido, Y., Dachlan, Y. P., and Kawamoto, F., 2010. Incidence and mutation analysis of glucose-6-phosphate dehydrogenase deficiency in Eastern Indonesian populations.*Acta. Med. Okayama.*, **64**(6), 367-373.

Turner, P.C., McLennan, A.G., Bates, A.D., and White, M. R. H. (2001). Molecular Biology. Second edition. New Delhi: BIOS Scientific Publishers Limited.

Vulliamy, T. J., Othman, A., Town, M., Nathwani, A., Falusi, A. G., Mason, P. J., and Luzatto, L., 1991. Polymorphic sites in the African population detected by sequence analysis of the glucose-6-phosphate dehydrogenase gene outline the evolution of the variants A and A-. *Proc. Nat. Acad. Sci.*, **88**, 8568-8571.

Vulliamy, T.J., Wanachiwanawin, W., Mason, .PJ.,and Luzzatto, L. 1989. G6PD Mahidol, a common deficient variant in South East Asia is caused by a (163)glycine-serine mutation. *Nucleic Acids Res.*, **17**, 5868.

Warrel, D. A., Cox, T. M., Firth, J. D., and Benz, E. J. (2005). Oxford: Oxford University Press.