Incidence and complete molecular characterization of glucose-6-phosphate dehydrogenase deficiency in the Guangxi Zhuang autonomous region of southern China: description of four novel mutations

Tizhen Yan*
Ren Cai*
OiuHua Mo
DongLin Zhu
Hong Ouyang
Lihua Huang
Mingguang Zhao
Fen Huang
Liyan Li
Xin Liang
Xiangmin Xu

Background and Objectives. Glucose-6-phosphate dehydrogenase (G6PD) deficiency is the most common human metabolic disorder in southern China. We investigated the incidence and distribution of mutations, the molecular pathology of affected females and the haplotype association with G6PD deficiency in patients from the Guangxi region.

Design and Methods. A population-based molecular analysis combining phenotypic screening and genotypic detection using both multiplex primer extension/denaturing high performance liquid chromatography assay and DNA sequence analysis were performed in a total of 4,704 individuals.

Results. The mutation frequency of male G6PD-deficient individuals was observed to be 7.43%. Twenty-seven genotypes from 361 individuals were found. Statistical analysis showed that there were significant differences in both the percentages of methemoglobin and the G6PD/6PGD ratio between heterozygote and hemizygote in males and between heterozygote and homozygote in females. However, no statistically significance was seen between hemizyotes and homozygotes. The mutation profile showed that five mutations, G6PD Kaiping^{1388A}, G6PD Canton^{1376T}, G6PD Gaohe^{95G}, Chinese-5^{1024T} and G6PD Viangchan^{871A}, are the most common in the area, accounting for 85% of the G6PD-deficient alleles. Ten rare mutations were detected in approximately 4% of the mutant chromosomes. Four novel mutations were found: G6PD Liuzhou^{442A}, G6PD Nanning^{703T}, G6PD Laibin^{1414C}, and G6PD Hechi$^{202A/871A}$. In addition, two other rare mutations, c.196T→A and c.202 G→A, were detected for the first time in Chinese patients. A single dominant haplotype (– – + – –) was observed in 94.0 % of 182 deficient chromosomes.

Interpretation and Conclusions. Our protocol could be used to extend the knowledge of molecular defects of G6PD gene in different geographical areas.

Key words: glucose-6-phosphate dehydrogenase (G6PD), G6PD gene, G6PD deficiency, missense mutation, neutral mutation, genotype.

Haematologica 2006; 91:1321-1328

*These authors contributed equally.

From the Department of Medical Genetics, Southern Medical University, Guangzhou 510515, Guangdong, P. R. China (TZ-Y, QH-M, H-OY, MG-Z, LY-L, XM-X); Liuzhou Municipal Maternity and Child Healthcare Hospital, Liuzhou 545001, Guangxi, P. R. China (R-C, D-LZ, L-HH, F-H, X-L).

Correspondence:
Xiang-Min Xu, Department of Medical Genetics, Southern Medical University, Tonghe 510515, Guangzhou, Guangdong, P.R.China.
E-mail: gzxuxm@pub.guangzhou.gd.cn

Glucose-6-phosphate dehydrogenase (G6PD; EC 1.1.1.49) deficiency is the most common human metabolic disorder, affecting an estimated 400 million people worldwide.[1,2] Although almost all affected individuals with G6PD-deficiency are generally not anemic and are asymptomatic, this defect may cause neonatal jaundice, mild hemolytic anemia to chronic non-spherocytic hemolytic anemia (CNSHA) with attacks of severe anemia induced by infections, specific drugs, or consumption of broad (fava) beans. Due to the fact that the *G6PD* gene is located on the X chromosome, the clinical symptoms of the disease are usually confined to hemizygous men, although female carriers with a marked expression of the aberrant allele may also suffer.[1] This deficiency in erythrocytes is usually found at high frequency in areas of the world where malaria is or has been endemic. Previous epidemiological studies showed that the geographic areas known to be *at risk* for G6PD deficiency include a broad region extending from the Mediterranean basin and parts of Africa, through the Middle East, the Indian subcontinent, Southeast Asia (SEA) including southern China, and parts of south America.[1,2,3] In the past decades there has been growing epidemiological confirmation of the hypothesis that G6PD deficiency protects against malaria mortality[4] and the application of novel haplotype-based techniques has demonstrated that malaria-protective genes have been subject to recent balancing selection.[5]

Over the past few decades, extensive studies have documented profiles of those G6PD variants exhibiting a wide range of phenotypes. Those affecting enzyme activity have been biochemically characterized.[1,6] To date, over 150 hematologically important mutations associated with different G6PD enzyme abnormalities, including the mutations responsible for CNSHA, have been identified at the molecular level in different ethnic populations of the world.[3,6] Molecular analysis has also revealed that each ethnic population has a characteristic profile of deficient variants. In southern China, where the G6PD deficiency was reported more

than 40 years ago, at least 18 different point mutations associated with reduced G6PD enzyme activity have been identified in the human G6PD gene.[3,7-12] Previous studies indicated a high population frequency of this defect in southern China, mainly distributed in the areas south of the Yangtze River. Population frequencies of G6PD deficiency in China ranged from 3.1%-16.1% in Guangdong,[8,9] 3.1%-9.0% in Taiwan,[10,11] up to 6.9% in parts of Sichuan,[8] and 16.7% among the Miao population.[12] However, the prevalence of G6PD deficiency in the Guangxi Zhuang Autonomous Region, which has a population of over 49 million, has not been determined. In addition, very little is known about the molecular basis of G6PD deficiency in female individuals of southern China because nearly all of the data on the incidence rate and mutation spectra were previously collected through patient-based investigations in male subjects only.[7-12]

In this report, we investigate the incidence and distribution of the deficiently mutations of the *G6PD* gene in this region of southern China using a population-based molecular analysis. The focus of this study was the molecular pathology of females with G6PD deficiency as well as haplotype-associations of common G6PD-deficient mutations.

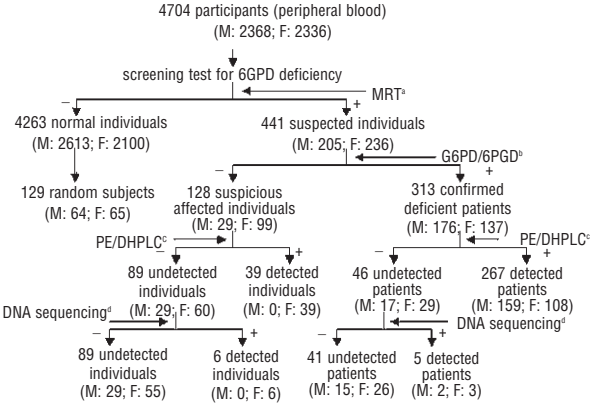

Figure 1. Diagnostic flow chart for the detection of G6PD deficiency in this study. The sample numbers detected at each step are indicated in brackets. [a]The negative and positive value of tests are >75% and ≤75%, respectively, for the MTR assay. [b]The negative and positive values of tests are >1.5 and ≤1.5, respectively, for G6PD/6PGD ratios. [c]The PE/DHPLC assay can detect ten G6PD mutations and a silent polymorphism: c.95 A→G, c.392 G→T, c.487 G→A, c.493 A→G, c.592 C→T, c.871 G>A, c.1024 C→T, c.1360 C→T, c.1376 G→T, c.1388 G→A, and c.1311 C→T. Three of them, c.487 G→A, c.493 A→G, and c.1360 C→T were not determined in our study. [d]For DNA sequencing, a positive result means the detection of the variants besides the above 11 known G6PD mutations or novel mutations. The 65 female samples randomly isolated from normal cohort were subjected to sequence analysis in order to characterize possible G6PD mutations in the heterozygote state in women in normal phenotypes. The eight mutations identified in this study were c.196T→A, c.202G→A, c.[202 G→A; 871 G→A], c.442 G→A, c.519C→T, c.703C→T, c.1004C→T and c.1414A→C (*Table 1*).

Design and Methods

Population

A total of 4704 unrelated participants comprising 2368 males (age 29.55±5.03 years) and 2336 females (age 27.55±3.89 years), whose registered parental origin is from 14 different Counties/Districts in Guangxi Zhuang Autonomous Region, were enrolled for this investigation between October 2002 and October 2003. Of these individuals, 98.5% were ethnic southern Chinese. Among them, most (94.8%) had Guangxi ancestry. Peripheral blood samples were randomly collected from these participants when they received their pre-marriage medical check-up at Liuzhou Municipal Maternity and Child Healthcare Hospital. For the molecular studies of G6PD variants to clarify their distribution and phenotype within this region, samples collected from the two major ethnic population groups, Han and Zhuang, which account for 64.3% (3027/4704) and 32.5% (1528/4704) of the total samples, respectively, were chosen. The remaining 149 individuals belonged to other minority groups. Informed consent to the investigation was obtained from all the participants.

Laboratory methods

The diagnostic flow chart for targeted detection of G6PD deficiency in this study is illustrated in Figure 1. Briefly, two different screening procedures, the conventional methemoglobin reduction test (MRT)[13] followed by the method advised by the WHO for measurement of G6PD/6PGD ratio,[14,15] were employed to measure the incidence of G6PD deficiency. The former was primarily used to estimate the prevalence of heterozygote females with G6PD deficiency in this region. In the second, genomic DNA was extracted by standard methods from all suspected positive samples and subjected to molecular analysis by a multiplex primer extension/denaturing high performance liquid chromatography (PE/DHPLC) assay, which can simultaneously genotype ten Southeast Asian G6PD deficiency-causing mutations and a silent polymorphism, as previously described.[16] Finally, direct DNA sequencing was performed on all the negative samples as determined by PE/DHPLC in order to detect any novel *G6PD* gene mutations. In this study, both coding exons and exon-intron boundaries in the human *G6PD* gene were completely sequenced using the ABI Prism DNA Analyzer (Model 377). As a control for this study, we randomly selected 129 samples from normal individuals and determined the G6PD mutation status by PE/DHPLC. Of these 129 samples, we sequenced all 65 female samples except for the two samples with known mutations detected by PE/DHPLC. The criteria used to make diagnosis of G6PD-deficiency in this study included: (i) a positive enzyme test by both MRT and G6PD/6PGD ratio assay or (ii) a positive mutation as determined by molecular analysis using PE/DHPLC or DNA sequencing. Blood counts and red-cell indices were assessed in all the samples using standard procedures.

Haplotype analysis

Five intragenic polymorphic sites at c.202 G→A (exon 4), c.376 A→G (exon 5), c.1116 A→G (exon 10), c.1311 C→T (exon 11), and c.1365-13T→C (intron 11) located in the *G6PD* gene, were used to construct haplotype patterns in the present study.[17-20] Haplotype analysis was conducted on the chromosomes from 176 male G6PD-deficient patients and six female individuals with newly identified mutations. To determine the frequency of 1311T in cis with intron 11-93 C in the normal population or in females, a double amplification refractory mutation system (ARMS) assay was designed for direct determination of the haplotype for these two polymorphic sites.

Statistical analysis

χ^2 analysis was used to assess statistically significant differences between the Han and Zhuang people. Completely randomized analysis of variance (ANOVA) was used to calculate statistically significant differences among hemizygote, homozygote, heterozygote, and normal control individuals. The percentages of methemoglobin and the G6PD/6PGD ratio of the three variants (c.1376 G→T, c.1388G→A and c.95 A→G) between the positive group (G6PD/6PGD≤1.5) and negative group (G6PD/6PGD>1.5) were analyzed by t-tests. Statistical analyses were conducted with an SPSS software program. The prevalences of different deficient G6PD alleles were calculated from a modified Hardy-Weinberg formula.

Results

Population prevalence and mutation spectrum of G6PD deficiency

Screened by MRT, 441 of the 4704 blood samples (9.38%) were suspected to be G6PD-deficient (positive). Of the 441 samples, 313 were classified as *confirmed G6PD-deficient patients* with a value ≤1.5 while the remaining 128 were clarified as *suspicious affected individuals* with a G6PD/6PGD ratio >1.5. These 128 samples underwent confirmatory testing. By molecular analysis using PE/DHPLC followed by direct DNA sequencing on samples from both these two groups as well as the normal population, we identified 272 (86.9% of 313 samples tested), 45 (35.2% of 128 samples tested) and 3 (2.33% of 129 normal samples tested) samples with a total of 15 mutant alleles in the *G6PD* gene. The number and percentage of different deficient chromosomes observed from our samples in both males and females are listed in Table 1. Overall, considering the numbers of mutant alleles contributed by 15 phenotypically positive males with unknown mutations, the total number of G6PD-deficient alleles in males was 176, given a frequency of 7.43% for G6PD deficiency in the Guangxi population. Comparison of the mutation frequency rates between the Han and Zhuang populations, 8.03% (122/1520) and 6.58% (50/760), respectively, demonstrated that there was no statistically significant difference (p=0.217) between ethnic backgrounds. The frequency of the common polymorphism 1311T in males is 7.8%, which is equivalent to the testing rate of 1311T in *cis* with intron 11-93 C.

Table 1. Results of screening and genotyping 4704 blood samples for G6PD deficiency in indiviudals in the Guangxi Zhuang Autonomous Region.

G6PD Variants	*G6PD Mutations*	*Males* *n*	*%*	*Females* *n*	*%*
Anant	c.1388 G→A	57	2.41%	66	1.41%
Canton	c.1376 G→T	51	2.15%	42	0.90%
Gaohe	c.95 A→G	32	1.35%	40	0.86%
Chinese-5	c.1024 C→T	10	0.42%	16	0.34%
Viangchan	c.871 G→A	8	0.34%	3	0.06%
Chinese-4	c.392 G→T	1	0.04%	2	0.04%
Fushan	c.1004C→T	-	-	3	0.06%
*Liuzhou	c.442 G→A	-	-	3	0.06%
Shunde	c.592C→T	-	-	1	0.02%
Miaoli	c.519C→T	-	-	1	0.02%
*Songklanagarind	c.196T→A	-	-	1	0.02%
*Asahi	c.202 G→A	-	-	1	0.02%
*Hechi	c.[202G→A;871 G→A]	1	0.04%	-	-
*Nanning	c.703C→T	1	0.04%	-	-
*Laibin	c.1414A→C	-	-	1	0.02%
	Unknown	15	0.63%	26	0.56%
Total		*176*	*7.43%*	*206*	*4.41%*

**New variants. The female subjects comprised 138 heterozygotes, 21 homozygotes and 26 with unknown mutations who were considered to be heterozygotes for calculation purposes.*

A total of 341 chromosomes with different known mutations were characterized in our study. About 85% of these mutations were accounted for by five common mutations (Table 2). The distribution of the five dominant mutant alleles (c.1388 G→A, c.1376 G→T, c.95 A→G, c.1024 C→T, and c.871 G→A) between the Han and Zhuang populations was not statistically different (p>0.1 for each of the five groups).

Contribution of the G6PD-deficient genotypes

We observed a total of 27 different genotypes from 317 G6PD-deficient individuals and three apparently *normal* individuals characterized through molecular analysis in this study, including 8 from hemizygotes, 13 from heteozygotes, and 6 from homozygotes, In addition, 15 males and 26 females had unknown mutations (Table 3). All of the eight male patients with a combination of c.871 G→A responsible for the variant G6PD Viangchan and the c.1311 C→T were found to have a *cis*-compound mutation which is similar to the linkage disequilibrium between mutation 871A and polymorphic site 1311 T previously observed in the Asian population.[17,21,22] As expected, most of genotypes in female homozygotes were common mutations, with 76.2% (16/21) of all homozygotes for c.1388 G→A mutation having four genotypes: c.[1388 G→A]+[1388 G→A], c.[1388 G→A]+[1376 G→T], c.[1388 G→A]+[95 A→G],

Table 2. The G6PD-deficiency chromosomes and their constitutions in Han and Zhuang people in Guangxi.

G6PD Variants	G6PD Mutations	Numbers of alleles Han n (%)	Zhuang n (%)	Other* n (%)	Overall n (%)
Anant	c.1388 G→A	81 (32.5)	38 (32.5)	4 (25.0)	123 (32.2)
Canton	c.1376 G→T	61 (24.5)	30 (25.6)	2 (12.5)	93 (24.3)
Gaohe	c.95 A>G	46 (18.5)	23 (19.7)	3 (18.8)	72 (18.8)
Chinese-5	c.1024 C→T	15 (6.02)	8 (6.84)	3 (18.8)	26 (6.81)
Viangchan	c.871 G→A	7 (2.81)	4 (3.42)	–	11 (2.88)
Chinese-4	c.392 G→T	2 (0.80)	–	1 (6.25)	3 (0.79)
Liuzhou	c.442 G→A	1 (0.40)	1 (0.85)	1 (6.25)	3 (0.79)
Fushan	c.1004 C→T	2 (0.80)	1 (0.85)	–	3 (0.79)
Asahi	c.202 G→A	1 (0.40)	–	–	1 (0.26)
Shunde	c.592 C→T	–	1 (0.85)	–	1 (0.26)
Miaoli	c.519 C→T	1 (0.40)	–	–	1 (0.26)
Songklanagarind	c.196 T→A	1 (0.40)	–	–	1 (0.26)
Nanning	c.703 C→T	1 (0.40)	–	–	1 (0.26)
Hechi	c.[202G→A ;871 G→A]	1 (0.40)	–	–	1 (0.26)
Laibin	c.1414A→C	1(0.40)	–	–	1(0.26)
Unknown		28 (11.2)	11 (9.40)	2 (12.5)	41 (10.7)
Total		249 (100)	117 (100)	16 (100)	382 (100)

**Sixteen G6PD-deficiency alleles were detected in 149 subjects belonging to other minority ethnic groups, including Melao, Tong, Yao, Hui, Buyi, Miao, Maonan, and Man. The c.1311 C→T silent polymorphism was not involved. The frequencies for the five most prevalent mutations between Han and Zhuang people were not statistically different according to χ^2 test analysis ($\chi^2 = 1.522$, $p>0.05$).*

and c.[1388 G→A]+[1024 C→T] (Table 3).

We then compared two phenotypes from female heterozygotes with the same genotypes who were isolated from the group of 39 women with initial suspicion of G6PD deficiency and the 108 women with confirmed G6PD deficiency in whom PE/DHPLC genotyping detected the mutation (Figure 1). The results showed a significant difference in phenotypic features of G6PD/6PGD measurements between the two groups and are compatible with the rule for incomplete inheritance of this X-linked inheritance, namely that the expression of deficiency varies as a result of variable X-inactivation in females. Thus, female heterozygotes may have variable enzyme activities even though they have the same mutation (Table 4). There were significant enzyme activity-related differences in both the percentages of methemoglobin and the G6PD/6PGD ratio between heterozygotes and hemizyotes ($p<0.001$), and heterozygotes and homozygotes ($p<0.001$). There was no statistically significant difference between hemizyote and homozygote ($p>0.1$) (Table 5).

Four novel mutations and two rare mutations

We identified four novel G6PD mutations in Chinese individuals. The first, designated G6PD Liuzhou[442A] (Genbank ID.DQ832762), involved a G→A substitution at nucleotide (nt) 442 within exon 5 of the *G6PD* gene, producing a Glu148Lys change in the protein. This novel mutation was detected in three female heterozygotes. The second mutation G6PD Nanning[703T] (Genbank ID.DQ 832761) involved a C→T change at nt 703 within exon 7 of the *G6PD* gene, resulting in a p.Leu235Phe substitution. The third, a variant identified in a male patient, involved a combination of c.202G→A responsible for the variant G6PD Asahi[23] and the common c.871 G→A (G6PD Viangchan) mutation. Since these two mutations have not previously been reported to exist in combination, we have called the new variant G6PD Hechi[202A/871A] (Genbank ID.DQ832766). The fourth, designated G6PD Laibin[1414C] (Genbank ID.DQ832763), the variant found in a normal female heterozygote, was an A to C substitution identified

Table 3. Twenty-seven different genotypes associated with G6PD deficiency and their contributions in Guangxi.

Hemizygote	n (%)	Heterozygote	n (%)	Homozygote	n (%)
c.1388 G→A	57 (32.4)	c.[1388 G→A]+[=]	46[a] (28.0)	c.[1376 G→T]+[1388 G→A]	5 (23.8)
c.1376 G→T	51* (29.0)	c.[1376 G→T]+[=]	35[b] (21.3)	c.[1388 G→A]+[1388 G→A]	4 (19.0)
c.95 A→G	32 (18.2)	c.[95 A→G]+[=]	31[c] (18.9)	c.[95 A→G]+[1388 G→A]	4 (19.0)
c.1024 C→T	10 (5.68)	c.[1024 C→T]+[=]	10[d] (6.10)	c.[95 A→G]+[1024 C→T]	3 (14.3)
c.871 G→A	8** (4.55)	c.[1004 C→T]+[=]	3[e] (1.83)	c.[1024 C→T]+[1388 G→A]	3 (14.3)
c.392 G→T	1 (0.56)	c.[871 G→A]+[=]	3[f] (1.83)	c.[95 A→G]+[1376 G→T]	2 (9.52)
c.703 C→T	1 (0.57)	c.[442 G→A]+[=]	3 (1.83)		
c.[202 G→A ;871G→A]	1 (0.57)	c.[392 G→T]+[=]	2 (1.22)		
		c.[196 T→A]+[=]	1 (0.61)		
		c.[202 G→A]+[=]	1 (0.61)		
		c.[519 C→T]+[=]	1 (0.61)		
		c.[592 C→T]+[=]	1 (0.61)		
		c.[1414A→C]+[=]	1 (0.61)		
Unknown	15† (8.52)		26†† (15.9)		
Total	176 (100)		164 (100)		21 (100)

**One c.1376 G>T case was linked with C→T mutation at nt 1311. **All individuals with G6PD Viangchan c.871G→A were in cis with c.1311C>T silent polymorphism. †Out of the 15 individuals, six subjects had the c.1311C→T silent polymorphism. ††Out of the 26 individuals, 11 subjects had the c.1311C→T silent polymorphism. [a]eight individuals with G6PD Kaiping c.1388 G>A had the c.1311C→T silent polymorphism. [b]four individuals with G6PD Canton c.1376 G→T had the c. 1311C→T silent polymorphism. [c]four individuals with G6PD Gaohe c. 95 A→G had the c.1311C→T silent polymorphism. [d]one individual with G6PD Chinese-5 c.1024 C→T had the c.1311C→T silent polymorphism. [e]All individuals with G6PD Fushan c.1004 C→T had the c.1311C>T silent polymorphism. [f]All individuals with G6PD Viangchan c.871G→A had the c.1311C→T silent polymorphism.*

Table 4. Comparison of phenotypic studies between two groups with the same mutations in female heterozygotes.

Phenotype Properties	c.1376 G→T Positive group (24)	c.1376 G→T Negative group (6)	c.1388 G→A Positive group (24)	c.1388 G→A Negative group (13)	c.95 A→G Positive group (17)	c.95 A→G Negative group (10)
G6PD/6PGD*	1.19±0.24	1.82±0.24	1.16±0.25	1.75±0.16	0.92±0.29	1.79±0.30
RBC ($\times10^{12}$/L)†	4.18±0.41	4.31±0.56	4.33±0.48	4.60±0.70	4.32±0.68	4.45±0.51
MetHb (%)*	29.16±25.02	55.88±10.88	29.74±25.41	52.58±22.42	20.65±31.11	43.14±29.60
MCV (fl)†	85.10±8.05	83.82±9.56	86.16±8.00	79.43±9.71	85.01±6.62	84.20±7.61
MCH (pg)†	31.03±11.35	28.67±3.38	29.19±3.24	26.60±3.46	32.84±13.05	28.91±3.27
Hb (g/L)†	125.71±14.81	123.83±17.84	123.04±29.35	117.58±11.67	130.12±17.56	128.40±7.56

*Values are given as mean±SD of (n) determinations. Each of two groups with the same mutations was compared by t- tests. *There were significant enzyme activity-related differences in both the percentages of methemoglobin and the G6PD/6PGD ratio of the three variants between the positive group (G6PD/6PGD≤1.5) and the negative group (G6PD/6PGD>1.5) (p<0.05). †No statistically significant difference in RBC, MCV, MCH, and Hb between the two groups (p>0.1). The normal reference values in adult females are as follows: G6PD/6PGD >1.5; RBC, red blood cells 4.0 –5.2×10^{12}/L; MetHb: methemoglobin >75%; MCH, mean corpuscular hemoglobin 26 ~34 pg; MCV, mean corpuscular volume 92±9 fl; Hb, hemoglobin 120-160 g/L. Note that the compound heterozygotes for these three variants and the 1311 polymorphism were not included in this comparative analysis to exclude the possible influence of the c.1311 C→T mutation.*

at nt 1414 in exon 12 of the *G6PD* gene, predicting a leucine 472 to isoleucine substitution. In addition to these, two other rare mutations were characterized in Chinese patients for the first time. One is G6PD Asahi (Genbank ID.DQ832764) existing on its own, which had been previously observed in a 3-year old Japanese boy with acute hemolysis;[23] the other is a T→A transversion at position 196, which was recently reported in Thailand and designated as a variant G6PD Songklanagarind[24] (Genbank ID.DQ832765). DNA sequence analysis of the *G6PD* gene from patients with the four novel mutations, as well as the variant G6PD Asahi mutation and variant G6PD Songklanagarind are presented in Figure 2. The data on the different mutations identified in the eight patients studied are summarized in Table 6. The new mutations reported here were found in the course of the population survey. We have been unable, despite repeated attempts, to obtain additional blood samples from these patients with newly identified mutations for further studies.

Haplotype association of common mutations

Haplotype patterns composed of five polymorphic sites in the 176 male G6PD-deficient subjects and six female subjects with newly identified mutations are shown in Table 7. We observed that most variants (94.0%), including the new variants G6PD Songklanagarind, G6PD Liuzhou^{442A}, G6PD Nanning^{703T}, and G6PD Laibin^{1414C}, share the same haplotype (– – + – –). The haplotype for the three G6PD variants corresponded to three other minor

Table 5. Relationship between the different genotypes and the G6PD/6PGD measurements.

Genotypes	Numbers of samples	MetHb (%) (M±SD)	G6PD/6PGD (M±SD)
Normal control*	129	91.63±4.57	1.70±0.30
Heterozygote	164	58.04±19.38	1.40±0.39
Homozygote	21	25.80±17.80	0.56±0.19
Hemizyote	176	23.31±18.97	0.45±0.20

*Comparisons of populations revealed statistically significant differences (p<0.05) by completely randomized ANOVA (129 G6PD normal individuals vs. the other three G6PD-deficient groups). There were significant enzyme activity-related differences in both the percentages of methemoglobin (MetHb) and the G6PD/6PGD ratio between the heterozygotes and hemizyotes (p<0.001), and between the heterozygotes and homozygotes (p<0.001). There was no statistically significant difference between hemizyotes and homozygotes (p> 0.1). *64 were males and 65 females. There was no significant difference between males and females from normal control (p=0.121, p>0.05).*

Table 6. Hematologic and genotypic data of four novel missense mutations and variants G6PD Songklanagarind and Asahi.

	Gender	Age	Hematological analysis MHb (%)	FST	G6PD/6PGD	NBT	Variant name	Molecular analysis cDNA nucleotide substitution	Amino acid substitution	Genbank ID.
Patient 1	F	27	60.5	+,++	2.21	–	Songklanagarind	c.196 T→A	p.F66I	DQ832765
Patient 2	F	25	29.3	±	0.91	+	Asahi	c.202 G→A	p.V68 M	DQ832764
Patient3	M	41	18.3	–	0.37	+	Hechi	c.[202 G→A; 871 G→A]	p.[V68M;V291M]	DQ832766
Patient 4	F	25	24.2	±	1.08	+	Liuzhou	c.442 G→A	p.E148K	DQ832762
Patient 5	F	30	74.0	+,++	1.71	+	Liuzhou	c.442 G→A	p.E148K	DQ832762
Patient 6	F	22	35.4	±	1.73	+	Liuzhou	c.442 G→A	p.E148K	DQ832762
Patient 7	M	44	57.8	±	1.33	+	Nanning	c.703 C→T	p.L235F	DQ832761
Patient 8	F	32	85.8	+,++	1.80	–	Laibin	c.1414A→C	p.I472L	DQ832763

FST: fluorescent Spot Test: complete absence of fluorescence (–),mild fluorescence(±), and bright fluorescence(+,++); NBT: nitroblue tetrazolium dye reduction test: normal (–), G6PD deficiency(+) . The normal reference values can be found in the legend of Table 4.

Table 7. Haplotype analysis of the G6PD variants in 176 males and six females with G6PD variants in five polymorphic sites.

G6PD mutations	*n*	*E4-202A NlaIII*	*E5-376G Fok I*	*E10-1116G PstI*	*E11-1311T BcI I*	*IVS-11-93C NlaIII*
c.1388 G→A	57	–	–	+	–	–
c.1376 G→T	50	–	–	+	–	–
c.95 A→G	32	–	–	+	+	+
c.1024 C→T	10	–	–	+	–	–
c.871 G→A	8	–	–	+	–	–
*c.442 G→A	3	–	–	+	+	+
c.[202 G→A ; 871 G→A]	1	+	–	+	–	–
c.392 G→T	1	–	–	+	–	+
c.703 C→T	1	–	–	+	–	–
*c.196 T→A	1	–	–	+	–	–
*c.202 G→A	1	+	–	+	–	–
*c.1414A→C	1	–	–	+	–	–
Unknown	15	–	–	+	–	–

+ indicates whether restriction site is present or whether c.1311 C→T mutation is present. n indicates the number of affected male subjects analyzed. "E" indicates the exon and "IVS" indicates the intron. The asterisk indicate six new G6PD mutations in females.

haplotypes. Eight subjects with variant G6PD Viangchan and one subject with variant Canton had haplotype (– – + + +). The two new variants G6PD Hechi[202A/871A] and G6PD Asahi have the haplotypes (+ – + – +) and (+ – + – –), respectively.

Discussion

We studied on the incidence and molecular characterization of G6PD deficiency in the Guangxi province located in southwestern China through screening and genotyping of blood samples from 4704 participants in this region. Guangxi is an autonomous region which has the largest population of minority ethnic groups in China, with over 60% of population belonging to the Han ethnic group and 32.6 % of the total population being Zhuang. The total population of 15.38 million accounts for 95% of the country's total. In this study, the prevalence of G6PD deficiency and the frequency of mutations were well demonstrated for two major ethnic groups, Han and Zhuang (Table 2). The results showed an unexpected correlation of both the rate and the mutation distribution between these two groups. The incidence and the distribution of G6PD alleles in Guangxi revealed by this study is reminiscent of the situation found with β-thalassemia in this region,[25] further confirming the hypothesis that both human *G6PD* and the β-globin gene are involved in malaria-protective selection factors. On average, the prevalence of G6PD deficiency was 7.43% (95% CI, 2.69 - 12.17) in the population of Guangxi. With a current annual birth rate of about 700,000 divided equally into males and females, we can estimate that the number of affected newborns in Guangxi at risk of G6PD-deficiency because of hemizygosity in males or homozygosity in females is 26,005 (95% CI, 19,500-45,000) and 1,932 (95% CI, 1,300- 3,680), respectively. The data obtained in our study, the high frequency of G6PD deficiency and the elucidation of mutation patterns enhances the basic knowledge of this disorder and also provides a more rational approach to population screening for prevention of and counselling on this disease in this area.

As illustrated in Figure 2, 15 mutations accounted for most of the G6PD variants identified in 161 male subjects and 159 female subjects. However, from the data compar-

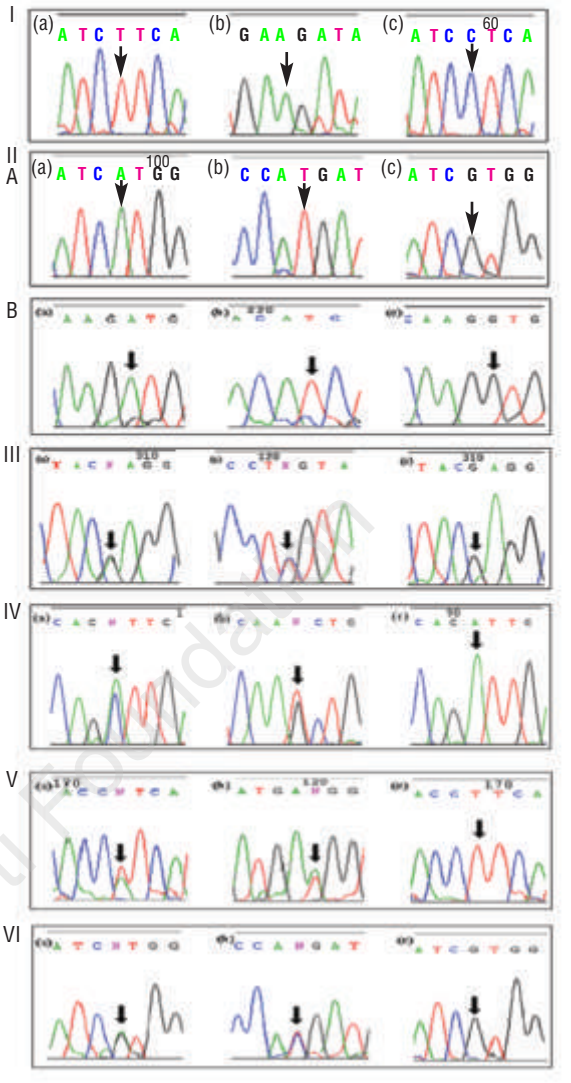

Figure 2. Partial nucleotide sequence of the four novel mutations, as well as the c.196 T→A and c.202 G→A mutations. A portion of DNA sequencing is presented for detection of c.703 C→T of G6PD Nanning in a male (I), c.[202 G→A;871 G→A] of G6PD Hechi in a male (II), c.442 G→A of G6PD Liuzhou in a female (III), c.1414A→C of G6PD Laibin in a female(IV), c.196 T→A of G6PD Songklanagarind in a female (V), and c.202 G→A of Asahi in a female (VI). For each of the sequencing results, the substitution base and nearby nucleotides are shown for the sense strand (a) the antisense strand (b) and wild-type control (c). The black arrow indicates the location of the mutation. The detection of G6PD Hechi double mutations is shown in panel II-A for c.202 G→A and II-B for c.871 G→A.

ing the disparity in the frequency of G6PD-deficient alleles between males (176/2368×100=7.43%) and females (206/(2336×2)×100=4.41%) obtained in our investigation (Table 1), the estimated misdiagnosis rates would be 40.7% due to the failure to test female heterozygotes using this protocol. Technically, the reason for the missed female heterozygotes may be due to insufficient sensitivity of the MRT assay, although this method has been determined to have the highest sensitivity for screening of G6PD deficiency in females.[26] However, when a majority of the screened positive samples and normal female samples were completely sequenced, only 9/200 exhibited novel mutations, including a G6PD Laibin^{1414C} heterozygote found in a phenotypically normal female. Unexpectedly, we detected two common mutations, the c.1388 G→A mutation and the c.1376 G→T mutation, in the normal female population (mutation frequency for each, 1.53%), suggesting that known G6PD mutations occur in female heterozygotes with a normal phenotype. Thus, we presume that the MRT assay did not pick up mutants including the known common G6PD variants (e.g. c.1388 G→A and c.1376 G→T) in the group of 2100 phenotypically normal female heterozygotes (Figure 1). Novel mutations underlying unknown G6PD variants may be: (i) a splicing mutation located at intron positions in the *G6PD* gene, where the sequences are missed by DNA sequence analysis; (ii) mutations involved in exonic splicing enhancer (ESE) sequences in the *G6PD* gene, which are currently described as new elements influencing the expression of genes;[27] (iii) mutations involved in the gene coding transcription factor at a locus on other chromosomes, which influence the expression of G6PD by a mechanism of trans-regulation, as probably occurs in β-thalassemia in humans caused by a mutation in an X-linked erythroid transcription factor.[28] Overall, the rational population-based protocol used in this study proved to be highly effective in detecting G6PD deficiency. It could be used to extend the knowledge of molecular defects of the *G6PD* gene in other geographical regions of China as well as in other countries. This protocol could be used to investigate the prevalence of *G6PD* deficiency genes in different geographical areas. We identified four novel G6PD-deficiently mutations, with a testing rate of 0.0213% for each of three alleles, G6PD Nanning^{703T}, G6PD Hechi$^{202A/871A}$, and G6PD Laibin^{1414C} and of 0.0638% for G6PD Liuzhou^{442A} allele in the local population. From the predicted enzyme activity, estimated using a structural model of human G6PD (SWISS-PhbViewer program), we can speculate that the p.E148K mutation may significantly influence the stability of the whole protein by changing the structure of the co-enzyme-binding domain (*data not shown*), suggesting that this variant could be able to cause severe clinical manifestation if it is in hemizygous state, in agreement with the findings in one of three female heterozygotes for this mutation (patient 4, Table 7) who had notable phenotypic manifestations. Comparison of the model structures for the wild type and the p.L235F or p.I472L mutant showed no major overall changes, suggesting that these two G6PD variants do not significantly decrease the stability of the protein compared to that of the wild type protein. We note that the male patient carrying the p.L235F mutant had a slightly decreased G6PD/6PGD ratio (1.33 in the mutant vs. >1.5 in the normal control). The G6PD variant Laibin could be regarded as a neutral mutation since the replacement of Leu472 by the structurally analogous amino acid Ile is generally unable to alter the protein function of G6PD. The G6PD c.[202G→A; 871G→A] mutation could have severe effects on enzyme activity because the ratio of G6PD/6PGD for the hemizygous patient with the double mutation was 0.37. It is easy to conceive that the *cis*-compound mutations in the *G6PD* gene could have cumulative effects. Accordingly, using software (*NNPREDICT*) prediction of the double mutations, four helix elements appended abnormally in the structural profile (*data not shown*) could destabilize the secondary structure of the protein, which may reduce the G6PD enzyme activity significantly. The data from haplotyping chromosomes from 176 male and 6 female G6PD-deficient individuals showed that the haplotype (– – + – –) is dominant, covering over 94% of the haplotype patterns in Chinese G6PD-deficient individuals, supporting the hypothesis that mutations occurring at nt 95, 1024, 1376, and 1388 could be derived from this ancient common haplotype. Two mutations, all 871 G→A mutations and very few 1376 G>T mutations, are linked with a minor haplotype (– – + + +), indicating that this mutation occurred more recently. Additional analysis of newly identified deficiency alleles in six females (three females with G6PD Liuzhou^{442A}, one with G6PD Laibin^{1414C}, one with G6PD Nanning^{703T} and one with G6PD Songklanagarind) displayed the same haplotype (– – + – –), further suggesting that they may have arisen from a common ancestral origin. The 871 G→A variant is prevalent in many parts of the world[3] but differs with respect to the 1311 polymorphism, a T existing in a population from Southeast Asia including south China described by the present study and other reports,[21] but a C in Indian[22] or Brazilian population,[29] suggesting that they may have arisen independently.

TZ-Y and R-C contributed equally to the work reported in this paper; XM-X conceived and designed the study; R-C, DL-Z, LH-H, F-H and X-L were responsible for clinical data collection and phenotypic analysis; TZ-Y, H-OY, MG-Z and LY-L conducted the molecular analyses; TZ-Y, R-C, QH-M and XM-X interpreted the data; TZ-Y and XM-X wrote the manuscript and created the tables and figures; TZ-Y, QH-M, H-OY and XM-X performed the statistical analysis. We thank all of our colleagues, past and present, who contributed to this project over the years. We thank Dr. Benjamin L. Legendre Jr. and Dr. Gary Lu for proofreading the manuscript and their excellent suggestions. his work was partially funded by a program grant from the National Natural Science Fund of China (NSFC) for Distinguished Young Scholars (30325037, to XM-X) and by a grant from the National Key Technologies R & D Program of China (2004BA720A04).

Manuscript received May 19, 2006. Accepted July 19, 2006.

References

1. Beutler E. G6PD: population genetics and clinical manifestations. Blood Rev 1996;10:45-52.
2. Miwa S, Fujii H. Molecular basis of erythroenzymopathies associated with hereditary hemolytic anemia: tabulation of mutant enzymes. Am J Hematol 1996;51:122-32.
3. Kotaka M, Gover S, Vandeputte-Rutten L, Au SW, Lan VM, Adams MJ. Structural studies of glucose-6-phosphate and NADP+ binding to human glucose-6-phosphate dehydrogenase. Acta Crystallogr D Biol Crystallogr 2005; 61:495-504.
4. Ruwende C, Khoo SC, Snow RW, Yates SNR, Kwiatkowski D, Gupta S. S, et al. Natural selection of hemi- and heterozygotes for G6PD deficiency in Africa by resistance to severe malaria. Nature 1995;376:246-9.
5. Verrelli BC, McDonald JH, Argyropoulos G, Destro-Bisol G, Froment A, Drousiotou A, et al. Tishkoff Evidence for balancing selection from nucleotide sequence analyses of human G6PD. Am J Hum Genet 2002;71:1112-28.
6. Johns Hopkins University. Online Mendelian Inheritance in Man. http://www.ncbi.nlm.nih.gov/entrez/dispomim.cgi?id=305900 (accessed March 2006)
7. Chiu DT, Zuo L, Chao L, Chen E, Louie E, Lubin B, et al. 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 1993; 81:2150-4.
8. Xu YK, Zheng RP, Liu LB, Hua XY, Zhong JP, Zhu XP, et al. Investigation of red cell glucose--6-phosphate dehydrogenase deficiency gene frequency in 9 nationality populations in 7 provinces (Autonomous Regions). Heredity Dis 1985;2:67-71.
9. Du CS, Xu YK, Hua XY, Wu QL, Liu LB, Glucose-6-phosphate dehydrogenase variants and their frequency in Guangdong, China. Hum Genet 1988; 80:385-8.
10. Chiang SH, Wu SJ, Wu KF, Hsiao KJ. Neonatal screening for glucose-6-phosphate dehydrogenase deficiency in Taiwan. Southeast Asian J Trop Med Public Health. 1999;30 Suppl 2: 72-4.
11. Tang TK, Huang WY, Tang CJ, Hsu M, Chen TA, Chen KH. Molecular basis of glucose-6-phosphate dehydrogenase (G6PD) deficiency in three Taiwan aboriginal tribes. Hum Genet 1995;95:630-2.
12. Du CS, Liu LB, Liu B, Tokunaga K, Omoto K. Glucose-6-phosphate dehydrogenase deficiency among three national minorities in Hainan Island, China. Gene Geogr 1988;2:71-4.
13. Brewer GJ, Tarlov AR, Alving AS. Methemoglobin reduction test - a new, simple in vitro test for identifying primaquine sensitivity. Bull World Health Organ 1960;22:633-40.
14. Du CS. Detection of Glucose-6-phosphate dehydrogenase deficiency in erythrocytes using method for measurement of G6PD/6PGD ratio. Birth Health Heredity 1991;4:1-3.
15. Beutler E, Blume KG, Kaplan JC et al: International Committee for Standardization in Haematology: recommended screening test for glucose-6-phosphate dehydrogenase (G6PD) deficiency. Br J Haematol 1979;43: 465-7.
16. Wu G, Liang WH, Zhu J, Ouyang H, Pearson PL, Mo QH, et al. Rapid, simultaneous genotyping of 10 Southeast Asian glucose-6-phosphate dehydrogenase deficiency-causing mutations and a silent polymorphism by multiplex primer extension/denaturing HPLC assay. Clin Chem 2005; 51: 1288-91.
17. Xu W, Westwood B, Bartscocas CS, Malcorra-Azpiazu JJ, Indrak K, Beutler E. Glucose-6 phosphate dehydrogenase mutations and haplotypes in various ethnic groups. Blood 1995; 85: 257-63.
18. Rodrigues MO, Freire AP, Pereira J, Martinis MD, Monteiro C. Glucose-6-phosphate dehydrogenase deficiency in Portugal: biochemical and mutational profiles, heterogeneity, and haplotype association. Blood Cells Mol Dis 2002;28:249-59.
19. Vulliamy T, Rovira A, Yusoff N, Colomer D, Luzzatto L, Vives-Corrons JL. Independent origin of single and double mutations in the human glucose 6-phosphate dehydrogenase gene. Hum Mutat 1996;8:311-8.
20. Kay AC, Kuhl W, Prchal J, Beutler E. The origin of glucose-6-phosphate-dehydrogenase (G6PD) polymorphisms in African-Americans. Am J Hum Genet 1992;50:394-8.
21. Tang TK, Liu TH, Tang CJ, Tam KB. Glucose-6-phosphate dehydrogenase (G6PD) mutations associated with F8C/G6PD haplotypes in Chinese. Blood 1995;85:3767-8.
22. Beutler E, Westwood B, Kuhl W. Definition of the mutations of G6PD Wayne, G6PD Viangchan, G6PD Jammu, and G6PD 'LeJeune'. Acta Haematol 1991;86:179-82.
23. Hirono A, Kawate K, Honda A, Fujii H. A single mutation 202G→A in the human glucose 6-phosphate dehydrogenase gene can cause acute hemolysis by itself Blood 2002;99:1498.
24. Laoombat V, Sattayasevana B, Janejindamai W, Viprakasit V, Shirakawa T, Nishiyama K, et al. Molecular heterogeneity of glucose-6-phosphate dehydrogenase (G6PD) variants in the south of Thailand and identification of a novel variant (G6PD Songklanagarind). Blood Cells Mol Dis 2005;34: 191-6.
25. Liang R, Liang S, Jiang NH, Wen XJ, Zhai JB, Nechtman JF, et al. α and β thalassaemia among Chinese children in Guangxi Province, P.R. China: molecular and haematological characterization. Br J Haematol 1994;86:351-4.
26. Zaffanello M, Rugolotto S, Zamboni G, Gaudino R, Tato L. Neonatal screening for glucose-6-phosphate dehydrogenase deficiency fails to detect heterozygote females. Eur J Epidemiol 2004;19:255-7.
27. Wu Y, Zhang Y, Zhang J. Distribution of exonic splicing enhancer elements in human genes. Genomics 2005; 86: 329-36.
28. Yu C, Niakan KK, Matsushita M, Stamatoyannopoulos G, Orkin SH, Raskind WH. X-linked thrombocytopenia with thalassemia from a mutation in the amino finger of GATA-1 affecting DNA binding rather than FOG-1 interaction. Blood 2002; 100: 2040-5.
29. Hamel AR, Cabral IR, Sales TS, Costa FF, Olalla Saad ST. Molecular heterogeneity of G6PD deficiency in an amazonian population and description of four new variants. Blood Cell Mol Dis 2002;28:399-406.