

Genetic Polymorphism of human CYP2D6 gene coding region from vivax malaria cases in Yunnan Province, China

Herong Huang

Dali University School of Basic Medicine

Ying Dong (✉ luxidongying@126.com)

Yanchun Xu

Yunnan Institute of Parasitic Diseases

Yan Deng

Yunnan Institute of Parasitic Diseases

Canglin Zhang

Yunnan Institute of Parasitic Diseases

Shuping Liu

Dali University School of Basic Medicine

Mengni Chen

Yunnan Institute of Parasitic Diseases

Yan Liu

Yunnan Institute of Parasitic Diseases

Research

Keywords: Plasmodium vivax, Relapse, Cytochrome P450 isoenzyme 2D6, Coding region, Primaquine, Radical Cure, enzyme activity, Phenotypes

Posted Date: August 1st, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-19379/v4>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background: Primaquine is one of two medications able to effectively eliminate the hypnozoites of *P. vivax*, and therefore has been recommended by World Health Organization (WHO) as the anti-relapsing drug for the treatment of vivax patients. However, 5-hydroxy-primaquine, the presumed active component, must be generated by CYP2D6 enzyme in human hepatocytes and reduced CYP2D6 enzyme activity caused by gene mutations has been considered cause of primaquine failure. Based on the analysis of CYP2D6 gene polymorphisms in suspected relapsed vivax malaria patients, the current study aims to determine the association between human CYP2D6 genotype and the efficacy of primaquine.

Methods: Blood samples from vivax patients in Yunnan Province who received "chloroquine/primaquine eight-day therapy" and experienced suspected relapses were collected from 2014 to 2018. Human genomic DNA was extracted, and 9 exons of CYP2D6 gene were amplified by polymerase chain reaction (PCR) and were then sequenced. The mutation types of CDS and their association with CYP2D6 enzyme activity changes were analyzed.

Results: A total of 156 blood samples from 75 relapsed cases of vivax malaria and 75 non-relapsed cases were collected and two nested-PCR products, the CYP2D6 gene fragments containing exon1-4 and exon5-9 (2411bp and 2388bp), were obtained from the sample. After splicing and combining these two amplification products, we found that the CDS of CYP2D6 gene in each sample was 1491bp in length. Moreover, we identified 24 haplotypes (Hap_1~Hap_24) in the Clustered CYP2D6 full-CDS of 150 patients, including 17 haplotypes in relapsed patients and 15 haplotypes in non-relapsed patients (8 haplotypes co-existed in two groups of patients). In addition, we identified 33 diplotypes (A_1~A_33) in 150 patients. The A_10 and A_12 diplotypes showed higher frequency in suspected relapsed group.

Conclusion: Among the numerous mutations of CYP2D6 gene, Hap_2 accounted for the largest proportion in 24 haplotypes. And the combined homozygous mutations at c. 408, c. 1457, especially the two loci combined mutations with c. 886, may be related to the poor efficacy of primaquine in the radical cure *P. vivax* in Yunnan, however, greater sample size is required confirm these findings.

Background

In recent years, the global malaria epidemic has gradually scaled down, yet the death tolls of malaria remain high, and the number of malaria cases in certain areas has rebounded [1]. Meanwhile, multiple technological limitations pose daunting challenges to the elimination of malaria, such as the difficulty in identifying low-density *Plasmodium* infection [2], the population expansion of drug resistant *Plasmodium* [3-4], and few new drugs in the antimalarial pipeline [5]. Although Chinese local cases of malaria (without a travel history to malaria epidemic areas outside China within the last 30 days) have not been reported since 2016 [6], China still needs to cope with a large number of imported malaria cases every year. Located in southwest China bordering Myanmar, Laos and other Southeast Asian countries (Fig.1), Yunnan Province has experienced large-scale malaria epidemic and was once the hardest-hit area of malaria [7]. Yunnan Province is still facing a great risk of imported malaria. And from 2011 to 2018, 50.7% of the imported malaria cases in Yunnan Province were caused by *P. vivax* [8]. Therefore, the elimination of malaria in Yunnan Province calls for the strict enforcement radical cure treatment to block the transmission of *P. vivax*.

Fig 1. The map of China highlighting the geographic location of Yunnan Province. Different colors represent different administrative areas, and green labels indicate the main source of infection in areas with imported cases.

Primaquine has proven to be effective in eradicating the hypnozoites of *P. vivax*, and the World Health Organization (WHO) recommends it for the treatment of relapsed malaria (and since 2018, the newly approved tafenoquine) [9]. The Cytochrome P450, family 2, subfamily D, polypeptide 6 (CYP2D6) enzyme in human hepatocyte metabolizes primaquine into 5-hydroxy-primaquine, the component thought to be responsible for killing *P. vivax* hypnozoites [10-12]. CYP2D6, also known as debrisoquine-4-hydroxylase, is an isozyme belonging to the Cytochrome P450 (CYP450) superfamily and an important phase I drug-metabolizing enzyme in the humans. CYP2D6 has low capacity and high affinity. Even though CYP2D6 only accounts for 4% of the total P450 enzyme protein in the liver, it is involved with the metabolism of ~30% of commonly used drugs, including primaquine [12]. Studies have found that the heterogeneity of CYP2D6 activity is predominantly governed by its genetic variation [13], and that decreased CYP2D6 isoenzyme activity caused by genetic polymorphisms would obstruct the generation of 5-hydroxy-primaquine, making large doses or repeated use of primaquine necessary to compensate for the declined efficacy of primaquine. Meanwhile, primaquine can cause life threatening hemolysis in patients with Glucose-6-Phosphate Dehydrogenase (G6PD) deficiency, subsequently inflicting dual hazards, including low efficacy of anti-relapse and acute hemolysis, on the patients [9,14-15]. So the patient's G6PD gene type should be identified by care workers prior to receiving anti-relapse treatment of vivax malaria and based on our and others' findings, it should be considered to test for CYP2D6 genotype and predicted enzyme activity as well [14,16-17].

CYP2D6 enzyme activity is measured mainly through phenotype observation and genotype analysis prediction. Compared with phenotype measurement, prediction method based on CYP2D6 genotype is characterized by the following advantages [18-19]. Firstly, genotypes do not change throughout one's life, and are not affected by environment or physical factors [19-20]. Secondly, genotyping only requires one specimen and can be completed before treatment [21]. In addition, using sequencing to determine the combination form of CYP2D6 alleles is a technique that can be easily obtained at present. Therefore, the use of genotyping method can create more stable results in predicting the phenotype of CYP2D6 enzyme activity [18,22-26]. The prediction is based on the locus status of CYP2D6 gene [13], or the allelic combination form of diplotype [27-28]. However, more than 113 CYP2D6 alleles have been defined by the Cytochrome P450 Nomenclature Committee at <https://www.pharmvar.org/gene/CYP2D6> [23]. The prediction of CYP2D6 enzyme activity by performing genotype analysis needs to be further optimized to enhance accuracy and systematicness [29].

Gene mutations of enzymes related to drug metabolism not only affect the clinical consequences of drug metabolism, but can even lead to individual adverse event [13]. Therefore, understanding the specific CYP2D6 genetic polymorphism in different individuals is the prerequisite to ensure radical treatment of vivax malaria based on personalized primaquine dose [24]. Despite the frequent use of primaquine in malaria endemic areas of Yunnan, no systematic survey about genetic polymorphisms of CYP2D6 has been conducted. Based on the investigation of CYP2D6 gene polymorphism, this study preliminarily explores the association between human CYP2D6 genotype and impaired efficacy of primaquine in the radical cure of *P. vivax* in order to identify population who may be refractory to primaquine treatment in Yunnan Province.

Materials And Methods

Subjects and blood sample

The subjects were diagnosed cases of *P. vivax* in Yunnan Province. Mono-infection of *P. vivax* was confirmed by both microscopy examination and *Plasmodium* 18S rRNA gene detection (PCR) by Yunnan Province Malaria Diagnosis Referent Laboratory (YPMDRL). All vivax malaria patients spent three days in the hospital and underwent three-days oral Chloroquine therapy (total 1550 mg), followed by an 8-day course of Primaquine (22.5 mg/day). And the eight days was directly observed therapy (DOT) by local Center for Disease Control (CDC) personnel to ensure compliance. The subjects were divided into two groups: (1) suspected relapsed group (SR group). The SR group consisted of patients who were re-diagnosed with vivax malaria within six months of follow-up after receiving above-mentioned treatment and who has not been exposed to malaria transmission again between two clinical malaria outbreaks through epidemiological investigation, and their infection was caused by the sequence similarity above 98% of *Plasmodium* in two samples detected by sequencing of *P. vivax* circumsporozoite protein (*pvcs*). A total of 75 SR cases were confirmed in Yunnan Province from 2014 to 2018. (2) Non-relapsed group (NR group). Seventy-five NR cases were selected by computer generated randomization from 143 cases of vivax malaria in 2018 who were followed up for one year and had no history of recurrence or signs. The infection sources of vivax cases were confirmed according to epidemiological investigation: those without a travel history to epidemic areas outside Yunnan Province within the last 30 days before the onset of malaria were defined as local cases; those have history of travelling to epidemic regions, such as Myanmar and Africa, were regarded as imported cases.

The blood samples of all the subjects were provided by the Centers for Disease Control and Prevention in Yunnan Province. Dried blood spots samples for genetic analysis were derived from 0.6 ml of venous blood, which were then put in a drying tube and stored at -80 °C before DNA extraction.

Extraction of human genomic DNA

Three dried blood spots (diameter = 5 mm) were made, and human genomic DNA was extracted by using QIAgen Mini Kit kit (Germany, QIAamp Company's DNA Mini Kit), according to the manufacturer's instruction. The extracted DNA was stored at -20 °C.

PCR amplification of CYP2D6 gene fragment

The primers of CYP2D6 for polymerase chain reaction (PCR) amplification were designed by using GenBank (<https://www.ncbi.nlm.nih.gov/gene/>), reference sequence (ID:NC_000022.11) was used as template, following the reaction condition in previous studies [23,30]. The primers for first-round nested PCR covering the coding region of CYP2D6 gene exons1~4 were: 5'-CCAGTGACAGATAAGGGTGC-3' and 5'-GACGTGGATAGGAGGTACAGAG-3'; the primers for second-round nested PCR were: 5'-GGTGACTTCTCCGACCAGG-3' and 5'-TTCCCAAACCCATCTATGC-3'. The final amplification region was 42131088 ~ 42128678, and the expected fragment length of the product was 2411 bp. The primers that cover exons 5~9 of CYP2D6 gene were: 5'-GCCGACTGAGCCCTGGGAGGTAGGTA-3' and 5'-GCTGGGGCCTGAGACTT-3. The amplification area was 42126035~42128422, and the amplification products had an expected fragment length of 2388 bp. For all the PCR reaction systems, we used 2.6µl DNA template, 14.0µl 2 × Taq PCR hybrid system (QIAGEN, Germany), 0.7µl upstream primer (20umol / L), 0.7µl downstream primers (20umol / L). The total volume was adjusted to 25.0µl with ddH₂O. PCR reaction conditions were clarified as follows: 92 to 94 ° C for 2-5 mins; 92 to 94 ° C for 10-30s, 50 to 56 ° C for 15-30s, 68 to 72 ° C for 2 to 3 min, 35 cycles; 68 to 72 ° C for 7 mins. The triplicated parallel repetition was adopted for each PCR reaction. The amplified products were observed using 1.5% agarose gel electrophoresis. The positive amplification products were sent to Shanghai Meiji Biomedical Technology Co, Ltd. for sequencing by using the dideoxy chain-termination method. Only the sequences that showed identical results in at least two tests were used for subsequent analysis.

Mutation analysis of CYP2D6 coding region

The sequencing results were aligned by using DNASTar v5.10 and BioEdit v7.0.9.0 software. The obtained all DNA sequences were assessed by using the Basic Local Alignment Search Tool (BLAST, <http://blast.ncbi.nlm.nih.gov/Blast.cgi>) in the NCBI platform. Sequences with identifications equals 100% and Query cover above 99% were considered as CYP2D6 gene sequence of human. The DNA sequence of exon 1~9 region of CYP2D6 gene was determined and spliced into CYP2D6 CDS in the order from exon1 to exon9. MEGA v5.04 software was used to align the CDS of each case with the wild-type sequence to confirm the loci of missense mutation and synonymous mutation. The defined mutation loci are queried for the corresponding ID in Genbank. DnaSP v6.11.01 software was used to identify missense mutation and synonymous mutation locus in each CDS, and the parameters of DNA diversity index (π) and expected heterozygosity (H_e) were calculated. Determine the allelic forms of the mutation loci by checking the sequencing peak map, the diplotypes of 150 cases were

determined and the frequency of each diplotype was calculated. SPSS software (version 21.0; IBN; Chicago; IL) was used to conduct chi-square test or Fisher's Exact Test on the frequency of haplotypes, diplotypes and mutation loci co-existed in SR and NR group, the inspection level was 0.05.

Results

Sample information and PCR amplification of human CYP2D6 gene

A total of 75 SR cases of *P. vivax* were enrolled, including 71 cases of one suspected relapsed event, 2 cases of two suspected relapse events, and two case of three suspected relapsed events. The geographic information of 75 NR cases is listed in Table 1. The male-female ratio was 2.66:1, and the proportions of cases imported from Myanmar, Africa countries, Laos and Yunnan were 96.7% (145/150), 1.3% (2/150), 0.7% (1/150) and 1.3% (2/150), respectively. The odds ratio of relapsed was calculated in different age groups and genders. Statistical significance was determined in age groups of 5-20 years (OR = 0.380; 95% CI: 0.986~0.146) and 21-60 years (OR = 2.471; 95% CI: 5.417~1.127). No significant correlation between other age groups or gender and the relapsed of vivax malaria was found.

Table 1. Information of 150 vivax malaria cases for amplification of CYP2D6 gene exon1-9 fragments

Variable	Total (n, F%)	SR group (n, F%)	NR group (n, F%)
Total	150(100)	75(50)	75(50)
Gender			
Male	109(72.7)	58(77.3)	51(68.0)
Female	41(27.3)	17(22.7)	24(32.0)
Age (in years)			
0-4	4(2.7)	3(4.0)	1(1.3)
5-20	23(15.3)	7(9.3)	16(21.4)
21-60	114(76.0)	63(84.0)	51(68.0)
above 60	9(6.0)	2(2.7)	7(9.3)
Malaria relapse			
1 episode	71(47.4)	71(94.6)	–
2 episodes	2(1.3)	2(2.7)	–
3 episodes	2(1.3)	2(2.7)	–
Infection source ^a			
Myanmar	145(96.7)	70(93.3)	75(100.0)
Africa	2(1.3)	2(2.7)	–
Laos	1(0.7)	1(1.3)	–
Yunnan indigenous	2(1.3)	2(2.7)	–

n: number of cases; F: Frequency; ^a Identified by epidemiological investigation

PCR amplification of two segments of human CYP2D6 gene, including exons1-4 and exons5-9, was performed by using 156 blood samples of relapsed cases and 75 blood samples of non-relapsed cases. The amplification products showed clear bands at 2000bp, which were considered as the target bands .

Polymorphism analysis of CYP2D6 gene coding region

Locus polymorphism of CDS chain

The PCR amplification products of CYP2D6 gene from the SR group and NR group were trimmed, and the CDS chains containing the complete exon1-9 (total length = 1491bp) were obtained from all samples. A total of 150 CDS chains (Genbank accession number: MT339044-MT339193) were selected from the 75 SR cases and 75 NR cases, and were then compared with wild-type sequence (NC: 000022.11). Base substitutions at 12 loci, such as c.31 and c.100, were determined (Table 2). All the mutation loci were found in Genbank and the corresponding ID, was regarded as result normal, and no new mutation loci were found. The proportions of third-base and first-base substitution in the codon triplet were 41.7% (5/12), and the proportion of second-base substitution was 16.6% (2/12). 7 missense mutation loci and 5 synonymous mutation loci were determined. Of these mutation loci, the SR cases accounted for 91.7% (11/12), whereas the mutation loci of NR cases accounted for 66.7% (8/12) (Table 2).

Comparison of the two groups differences in mutation allele frequency at the 12 SNPs (Table 2). In the SR group, the mutation allele frequency was the highest in c.1457 G>C (80.7% $\frac{121}{150}$), followed by c. 408 G>C (73.3% $\frac{110}{150}$). In the NR group, on the contrary, the mutation allele frequency was the highest in c. 408 G>C (80.0% $\frac{120}{150}$), followed by c. 1457 G>C (74.7% $\frac{112}{150}$). Regarding the two mutation alleles, a non-significant trend ($\chi^2=1.863$ and $\chi^2=1.557$, $P > 0.05$) for difference in frequency distribution among groups SR and NR was detected. However, the frequency differences of the other two mutation alleles c. 100C>T and c. 886C>T between the two groups were statistically significant. The variant c. 100C>T allele showed lower frequency in SR group ($\chi^2=9.039$, $P < 0.05$). But the variant c. 886C>T allele showed higher frequency in SR group ($\chi^2=8.165$, $P < 0.05$).

Table 2. Polymorphism and mutation allele frequency comparison of SR group and NR group in the coding domain of CYP2D6 Genes from 1st aa to 497th aa

SR group				NR group				p ^c	SNP ID in Genbank
Loci	Codon change ^a	Amino acid change	No. (n=150% ^b)	Loci	Codon change ^a	Amino acid change	No. (n=150% ^b)		
c.31	GTG ATG	V11M	1(0.7)	–	–	–	–	–	rs769258
c.100	CCA TCA	P34S	66(44.0)	c.100	CCA TCA	P34S	92(61.3)	0.003 ^d (S)	rs1065852
c.271	CTG ATG	L91M	4(2.7)	c.271	CTG TTG	L91L	2(1.3)	0.680 ^d (NS)	rs28371703
c.281	CAC CGC	H94R	4(2.7)	–	–	–	–	–	rs28371704
c.294	ACC ACG	T98T	4(2.7)	c.294	ACC ACG	T98T	1(0.7)	0.367 ^d (NS)	rs28371705
c.297	GCC GCT	A99A	1(0.7)	–	–	–	–	–	rs200269944
c.336	TTC TTT	F112F	61(40.7)	c.336	TTC TTT	F112F	76(50.7)	0.082 ^d (NS)	rs1081003
c.408	GTG GTC	V136V	110(73.3)	c.408	GTG GTC	V136V	120(80.0)	0.172 ^d (NS)	rs1058164
c.505	GGT AGT	G169S	2(1.3)	–	–	–	–	–	rs5030865
–	–	–	–	c.801	CCC CCA	P267P	2(1.3)	–	rs28371718
c.886	CGC TGC	R296C	29(19.3)	c.886	CGC TGC	R296C	12(8.0)	0.004 ^d (S)	rs16947
c.1457	AGC ACC	S486T	121(80.7)	c.1457	AGC ACC	S486T	112(74.7)	0.212 ^d (NS)	rs1135840

^a DNA base highlighted in bold indicates the occurrence of SNP; n: number of chromosomes; b: Detection frequency, denominator is 150; ^c Analysis of mutation

allele frequency difference between two groups; ^d Chi-square test; NS: not significant; S: significant (P<0.05).

Polymorphism of haplotypes

A total of 24 CYP2D6 haplotypes (Hap_1–Hap_24) were identified amongst the 150 vivax malaria samples; 17 haplotypes were observed in the sequences of SR group ($\pi=0.0015$, $He= 0.8191$). 15 haplotypes were observed in the sequences of NR group ($\pi = 0.0014$ and $He = 0.8065$). 8 haplotypes, including Hap_2, Hap_3, Hap_4, Hap_5, Hap_6, Hap_7, Hap_14 and Hap_17, were found in both SR and NR group. Haplotype and frequency were shown in Fig. 1. Hap_3 was wild-type sequence, and the rest were mutant sequence. Hap_2 accounted for the largest proportion of 36.6% (55/150), followed by 15.3% for Hap_3 (23/150), 8.6% for Hap_4 (13/150), 6.6% for Hap_5 (10/150), 8.0% for Hap_6 (12/150), 6.6% for Hap_7 (10/150) and 4.6% for Hap_21 (7/150). The proportions of Hap_1, Hap_8, Hap_9, Hap_10, Hap_12, Hap_13, Hap_15, Hap_16, Hap_18, Hap_19, Hap_20, Hap_22, Hap_23, and Hap_24 were as low as 0.7% (1/150). The proportions of Hap_11, Hap_14 and Hap_17 haplotypes were 1.3% (2/150).

Fig 1. 9 exons are indicated by numbered boxes with DNA polymorphisms indicated on top. Predicted amino acid changes are indicated below; no change is synonymous mutation. SR: Suspected relapsed group; NR: Non-relapsed group.

Polymorphism of diplotypes

The results of allelic form recognition to mutation loci showed that the CYP2D6 gene in 150 cases could be defined as 33 diplotypes (Table 3). Among them, only one diplotype (A₁) was non-mutation homozygous in 12 SNPs, the other 32 diplotypes showed mutant heterozygous at least one locus (A₂₁). In the 32 mutant diplotypes, the allelic form of 96 mutation loci accumulated in 12 loci were mutant homozygous accounted for 72.9% (70/96). Although there was no regularity in the number of mutation loci between the SR group and the NR group, the diplotype A₁₇ showed mutations at as many as 7 loci (c. 100, c. 271, c. 281, c. 294, c. 297, c. 408 and c. 1457), and the diplotype came from a patient suspected of relapsed vivax malaria for 3 times. In addition, The diplotypes A₁₀ and A₁₂ showed mutant homozygous at c.408, c.1457 and c.408, c.886, c.1457, respectively, and that the frequency of SR group was significantly higher than that of NR group (Fisher's Exact Test $P < 0.05$) (Table 3). But the frequency of diplotype A₃₀ in the NR group was significantly higher than that in the SR group (Fisher's Exact Test $P < 0.05$) (Table 3).

Table 3 Analysis the diplotypes in SR group and NR group

Haplotypes	Diplotypes	Mutation loci and allelic form ^a												Case groups	
		c.31	c.100	c.271	c.281	c.294	c.297	c.336	c.408	c.505	c.801	c.886	c.1457	SR group	NS (F)
NC_000022.11	-	G/G	C/C	C/C	A/A	C/C	C/C	C/C	G/G	G/G	C/C	C/C	G/G	-	-
Hap_3	A_1	G/G	C/C	C/C	A/A	C/C	C/C	C/C	G/G	G/G	C/C	C/C	G/G	8(10.7)	1
Hap_1	A_2	G/G	C/C	C/C	A/A	C/C	C/C	C/C	C/C	G/A	C/C	T/T	C/C	1(1.3)	0
Hap_2	A_3	G/G	T/T	C/C	A/A	C/C	C/C	T/T	C/C	C/C	C/C	C/C	C/C	27(36.0)	2
	A_4	G/G	G/T	C/C	A/A	C/C	C/C	T/T	C/C	C/C	C/C	C/C	C/C	0	1
	A_5	G/G	G/T	C/C	A/A	C/C	C/C	T/T	G/C	C/C	C/C	C/C	G/C	0	1
Hap_4	A_6	G/G	T/T	C/C	A/A	C/C	C/C	T/T	C/C	C/C	C/C	C/C	G/C	1(1.3)	1
	A_7	G/G	T/T	C/C	A/A	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	0	2
	A_8	G/G	G/T	C/C	A/A	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	2(2.7)	6
Hap_5	A_9	G/G	G/T	C/C	A/A	C/C	C/C	C/C	C/C	C/C	C/C	C/C	G/C	0	3
	A_10	G/G	C/C	C/C	A/A	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	8(10.7)	1
	A_11	G/G	C/C	C/C	A/A	C/C	C/C	C/C	C/C	C/C	C/C	C/C	G/C	0	1
Hap_6	A_12	G/G	C/C	C/C	A/A	C/C	C/C	C/C	C/C	C/C	C/C	T/T	C/C	8(10.7)	0
	A_13	G/G	C/C	C/C	A/A	C/C	C/C	C/C	C/C	C/C	C/C	G/T	C/C	2(2.7)	2
Hap_7	A_14	G/G	C/C	C/C	A/A	C/C	C/C	4(5.4)	1						
	A_15	G/G	C/C	C/C	A/A	C/C	G/C	3(4.0)	2						
Hap_8	A_16	G/G	C/C	C/C	A/A	C/C	C/C	T/T	C/C	C/C	C/C	T/T	C/C	1(1.3)	0
Hap_9	A_17	G/G	T/T	A/A	G/G	G/G	G/T	C/C	C/C	C/C	C/C	C/C	C/C	1(1.3)	0
Hap_10	A_18	G/G	C/C	C/C	A/A	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	1(1.3)	0
Hap_11	A_19	G/G	C/C	C/C	A/A	C/C	C/C	C/C	C/C	C/C	C/C	T/T	G/C	2(2.7)	0
Hap_12	A_20	G/G	T/T	A/A	G/G	G/G	C/C	C/C	C/C	C/C	C/C	C/C	C/C	1(1.3)	0
Hap_13	A_21	G/G	T/T	C/C	A/A	C/C	C/C	T/T	C/C	C/C	C/C	T/T	C/C	1(1.3)	0
Hap_14	A_22	G/G	C/C	C/C	A/A	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	1(1.3)	1
Hap_15	A_23	G/G	C/C	C/C	A/A	C/C	C/C	C/C	C/C	C/C	C/C	G/T	C/C	1(1.3)	0
Hap_16	A_24	G/A	C/C	C/C	A/A	C/C	G/C	1(1.3)	0						
Hap_17	A_25	G/G	G/T	C/C	A/A	C/C	C/C	T/T	C/C	C/C	C/C	C/C	C/C	0	1
	A_26	G/G	T/T	C/C	A/A	C/C	C/C	G/T	C/C	C/C	C/C	C/C	C/C	1(1.3)	0
Hap_18	A_27	G/G	T/T	C/C	A/A	C/C	C/C	T/T	C/C	C/C	C/C	C/C	C/C	0	1
Hap_19	A_28	G/G	G/T	C/C	A/A	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	0	1
Hap_20	A_29	G/G	C/C	C/C	A/A	C/C	C/C	C/C	C/C	C/C	A/A	T/T	C/C	0	1
Hap_21	A_30	G/G	T/T	C/C	A/A	C/C	C/C	C/C	C/C	C/C	C/C	G/T	C/C	0	7
Hap_22	A_31	G/G	C/C	T/T	A/A	G/G	C/C	T/T	C/C	C/C	C/C	C/C	C/C	0	1
Hap_23	A_32	G/G	C/C	C/C	A/A	C/C	C/C	T/T	G/C	C/C	C/C	G/T	G/C	0	1
Hap_24	A_33	G/G	C/C	C/C	A/A	C/C	C/C	T/T	C/C	C/C	C/C	C/C	C/C	0	1

^a DNA base highlighted in bold indicates the occurrence of SNP; non-mutation is indicated by lines; n: number of cases; F: Frequency; ^bAnalysis of haplotype f between two groups; ^cChi-square test; ^dFisher's Exact Test; NS: not significant; S: significant (P<0.05).

Discussion

Located on chromosome 22q13.1, CYP2D6 gene (4383bp in length) consists of 8~9 exons and 7~8 introns. Its CDS is 1338~1491bp in length and encodes 446~497 amino acids. The heterogeneity of CYP2D6 monooxygenase activity varies from complete dysfunction to ultra-rapid enzyme metabolism, depending on the various genotypes of CYP2D6 [14]. Among the 150 identified alleles, multiple mutations such as CYP2D6*3, CYP2D6*4, CYP2D6*5 and CYP2D6*6 are considered to contribute to compromised CYP2D6 enzyme activity [32,33].

To our knowledge, the current study is the first one that attempts to analyze mutation polymorphism in the whole coding region of human CYP2D6 gene, thereby verifying its possible association with thwarted CYP2D6 enzyme activity in relapsed vivax malaria patients after primaquine treatment. Among the 24 haplotypes defined, Hap_2 makes up the largest proportion of the sequences, reaching as high as 36.6% (55/150). This haplotype shows multiple mutations at four loci, including c.100C > T, c.336C > T, c.408G > C, c.1457G > C. These mutations are different from previous studies which reported that CYP2D6*10 allele is frequently detected in Asian populations [32, 34]. CYP2D6*10 allele has multiple mutations at c. 100(C > T), c. 408(G > C) and c. 1457(G > C), yet is devoid of c. 336C > T mutation as opposed to Hap_2. Such disparity might be attributed to fact that the sequencing of the whole gene coding region can easily detect all base substitutions. However, the factor of different population should be taken into account to determine whether such differences in common alleles of CYP2D6 gene are worthy of in-depth research. The population studied in this paper is not randomly selected, but is composed of vivax malaria patients residing in endemic areas, and 50% of them (75/150) had relapsed malaria. Thusly, the genotypes of such population may show the accurately selective outcome of human defective genes after the screening of malaria epidemic.

Although decreased enzyme activity in the heterogeneity of CYP2D6 is speculated to lead to failed therapy of primaquine to kill hypnozoite of *P. vivax* [9,17,35-37], the correlation between different genotypes and the relapse of *P. vivax* is poorly understood. In this study, we found that the genotypes with combined mutations at c. 408, c. 886 and c. 1457 were most closely related to the relapse of vivax malaria (OR = 5.615, P < 0.05) (Table 3), and that mutation at c. 886 locus is the key indicator that determines vivax malaria relapse. Such findings are not consistent with that of Silvino et al [37], which believes that c. 100, c. 321 and other locus mutations are more closely related to vivax malaria relapse. Such disparity might be attributed to the different methods used for detecting gene mutation loci. Silvino et al. employed restriction enzyme digestion analysis to identify individual mutation locus, such as c.100 and c.321. Nevertheless, we applied whole gene coding region sequencing method to find locus polymorphism. Of note, in determining the close association between multiple mutations at c. 408, c. 886 and c. 1457 with the relapse of vivax malaria, we found that intron c.1023+36 mutation may play a critical role in aggravating damaged CYP2D6 enzyme activity. In the 4 relapsed cases of c. 886 mutation heterozygosity (C/T) or homozygous mutation (T/T), mutations were found at c. 1023+36, which was only the mutant heterozygotes. This finding concurs with c. 1023+36 mutations found by Silvino et al [37] and Raimundo et al [38] in patients suffering from failed treatment of vivax malaria. Additionally, the multiple mutation genotypes at c.100, c.408 and c.1457 loci determined in the preset study depicted a unique phenomenon of reduced risk of vivax malaria relapse (OR = 0.159, P < 0.05), and such findings warrants further analysis and confirmation.

To date, direct prediction method based on CYP2D6 locus mutation polymorphism, which was summarized by Zanger et al [27] and indirect prediction method of genotype scoring, which was recommended by Gaedigk et al [28] and Caudle et al [31], are applied to predict CYP2D6 enzyme activity based on genetic information. Since the genotype scoring method can quantify the degree of gene mutation and abides by more abundant genotypic criteria, this method was used in the study. However, no matter Hap_6 indicates high risk of relapse or Hap_4 functions as the protective factors, we predicted that all the genotypes are above IM level. CYP2D6*2 is the allele type that indicates highest risk of relapse, and most of the non-relapsed patients had CYP2D6*10 allele. This result is not in line with that of Brasil et al [36], which reported that most of the alleles in relapsed patients are CYP2D6*10, CYP2D6*39 and so forth. Moreover, it contradicts the clinical observation of relapsed patients in this study. Such conflict of outcome might be justified that the prediction method of CYP2D6 enzyme activity used in this study focuses on analyzing the whole gene coding region information, thereby taking more alleles into account and leading to enhanced genotype score. In contrast, Brasil et al., detected only a few mutation loci and therefore could miss out the reverse effects of other locus mutations. CYP2D6*2 allele shows heterogeneous phenotypes exhibiting both normal and abnormal enzyme activity [39]. Therefore, identifying the heterogeneous CYP2D6 with similar genotypes is the key to predicate the enzyme activity of CYP2D6 [20, 28], regardless of the method used for detection. Meanwhile, the prediction method should be further optimized, because the effect of primaquine resistance in patients with relapsed vivax malaria is not eliminated.

Although this study applied first-generation sequencing method to analyze polymorphism in the coding region of CYP2D6 gene, the strategy of triplicated parallel repetition was adopted for each PCR reaction and the sequencings of its products. Such effort helped eliminate the rate of error in single PCR sequencing to a certain extent. Meanwhile, the cost-effective sequencing analysis method is more suitable for molecular epidemiological cross-sectional investigation of CYP2D6 gene mutation in the endemic areas where primaquine is used to intervene vivax malaria. However, this study has some limitations. Firstly, the sample size is not large enough, and the low-frequency alleles of CYP2D6 genes detected in the study should be further verified. Secondly, all the non-relapsed patients were selected from 2018, which is not equivalent to the relapsed patients enrolled for four consecutive years. Sampling bias may affect the experimental results due to changes of various factors over the years. In addition, other confounding factors which may amplify the identification of relapsed vivax malaria patients, such as the existence of primaquine-resisted *Plasmodium*, are not ruled out, despite the identification of relapsed cases has been achieved by comparing the genetic origin of infected strains. Finally, CYP2D6 enzyme activity was analyzed by assigning the activity score of zygotic allelic, a method recommends by Caudle et al [31], yet this study defined more CYP2D6 gene locus polymorphisms. Since there is no benchmark standard to

scores the activity of these loci polymorphisms, the prediction of CYP2D6 enzyme activity in this study may not be completely homogenous with the findings of Caudle et al.

Conclusion

In conclusion, the prediction of CYP2D6 enzyme activity based on whole gene coding region sequencing preliminarily reveals CYP2D6 gene mutation polymorphisms in relapsed *P. vivax* patients who showed poor curative effect of primaquine in Yunnan. This study makes up for the analysis of genetic correlation factors in relapsed *P. vivax* patients in China, and reaches the following findings: (1) Among the multiple mutations of CYP2D6 gene, combined mutations at three loci, including c.408, c.886 and c.1457, are most closely related to decreased CYP2D6 enzyme activity. Mutations at c.408 and c.1457 induce compromised CYP2D6 enzyme activity, yet a larger sample size is need to validate whether such decreased activity is aggravated by c.886 mutation; (2) The key to predict CYP2D6 enzyme activity is to define the genotype accurately and accordingly, based on the genetic correlation between genotype and enzyme activity phenotype with reasonable sample size. Because the whole gene sequencing method can fully identify the mutation loci, it is a promising molecular epidemiological method for investigating CYP2D6 gene mutation polymorphism.

Abbreviations

WHO: World Health Organization; CYP2D6: Cytochrome P450, family 2, subfamily D, polypeptide 6; G6PD: Glucose-6-Phosphate Dehydrogenase; CDS: coding DNA sequence; PCR: polymerase chain reaction; PM: Poor metabolizer; IM: Intermediate metabolizer; NM: Normal metabolizer; UM: Ultra-rapid metabolizer; *pvcsp*: *Plasmodium vivax* circumsporozoite protein; Hap: haplotype; π : Diversity index; He: Expected heterozygosity; OR: odds ratio;

Declarations

Acknowledgements

We appreciate the support from the Centers for Disease Control and Prevention in states/cities and counties such as Dehong, Baoshan, Kunming, Pu'er, Lincang, Dali, Nujiang, Lijiang, Xishuangbanna, Yuxi, Chuxiong, Honghe, Zhaotong, Diqing, Qujing, and Whenshan.

Authors' contributions

HH was responsible for the gene testing, study design, statistics and analysis of the data, and wrote the manuscript; YDong performed coordination of all project, study design, analysis of the data and the manuscript guidance and revision. SL participated in the study design; YX, YL, YDeng, MC, CZ performed the collection blood samples and microscopy examination. All authors read and approved the final draft of the manuscript.

Funding

Supported by National Science Foundation (No. 81660559,81960579), the Youth Project of Yunnan Province Basic Research Program (No. 2017FD007), and Research institutions in medical and health units of Yunnan Province Research projects No. 2018NS0180

Availability of data and materials

Not applicable.

Ethics approval and consent to participate

The study was approved by Yunnan Institute of Parasitic Diseases and by the Ethical Committee. Genetic testing experiment, etc. were performed on stored blood samples obtained as part of routine diagnostic work-up patients with fever suspected of malaria. Although the absence of risk and the anonymous data processing are ensured among the study, consent from potential malaria patients need to be obtained during collecting blood samples. Demographic, clinical and epidemiological information of each fever patient will be collected. Database access will be restricted by password, and Yunnan Institute Parasitic Diseases will not allow retrieving and saving the personal identification information into the project database. It is committed not to provide information about the patient to any person unrelated to the study.

Consent for publication

All authors provided their consent for the publication of this report.

Competing interests

The authors declare that they have no competing interests.

Author details

¹School of Basic Medical Sciences, Dali University, Dali 667000, China. ²Yunnan Institute of Parasitic Diseases Control, Yunnan Provincial Key Laboratory of Vector-borne Diseases Control and Research, Yunnan Centre of Malaria Research, Academician Workstation of Professor JinNingyi, Pu'er Academician Workstation of Professor Lubin Jiang, Pu'er, 665000, China

References

1. World malaria report 2018. Geneva: World Health Organization; 2018. <https://www.who.int/malaria/publications/world-malaria-report-2018/en/>. Accessed 10 Mar 2020.
2. Recker M, Bull PC. Recent advances in the molecular epidemiology of clinical malaria. *F1000Res*. 2018; 7.
3. Thuy-Nhien N, Tuyen NK, Tong NT, Vy NT, Thanh NV, Van HT, et al. K13 propeller mutations in *Plasmodium falciparum* populations in regions of malaria endemicity in Vietnam from 2009 to 2016. *Antimicrob Agents Chemother*. 2017; 61.
4. Dong Y, Wang J, Sun A, Deng Y, Chen M, Xu Y, et al. Genetic association between the Pfk13 gene mutation and artemisinin resistance phenotype in *Plasmodium falciparum* isolates from Yunnan Province, China. *Malar J*. 2018; 17:478.
5. WHO Malaria Policy Advisory Committee and Secretarial. Malaria Policy Advisory Committee to the WHO: conclusions and recommendations of September 2012 meeting. *Malar J*. 2012; 11:424.
6. Feng J, Zhang L, Huang F, Yin ZH, Tu H, Xia ZG, et al. Ready for malaria elimination: zero indigenous case reported in the People's Republic of China. *Malar J*. 2018; 17:315.
7. Feng J, Xiao H, Zhang L, Yan H, Feng X, Fang W, et al. The *Plasmodium vivax* in China: decreased in local cases but increased imported cases from Southeast Asia and Africa. *Sci Rep*. 2015; 5:8847. Published 2015 Mar 5.
8. Zhang L, Feng J, Tu H, Xia ZG, Zhou SS. Challenges in malaria elimination the epidemiological characteristics of *Plasmodium vivax* in China from 2011 to 2018. *Chin J Parasitol Parasit Dis*. 2019; 37(05):532-538.
9. Bennett JW, Pybus BS, Yadava A, Tosh D, Sousa JC, McCarthy WF, et al. Primaquine failure and cytochrome P-450 2D6 in *Plasmodium vivax* N Engl J Med. 2013; 369: 1381-2.
10. Marcsisin SR, Reichard G, Pybus BS. Primaquine pharmacology in the context of CYP2D6 pharmacogenomics: Current state of the art. *Pharmacol Ther*. 2016; 161:1-10.
11. Vale N, Moreira R, Gomes P. Primaquine revisited six decades after its discovery. *Eur J Med Chem*. 2009; 44:937-53.
12. Pybus BS, Sousa JC, Jin X, Ferguson JA, Christian RE, Barnhart R, et al. CYP450 phenotyping and accurate mass identification of metabolites of the 8-aminoquinoline, anti-malarial drug primaquine. *Malar J*. 2012; 11:259.
13. Twist GP, Gaedigk A, Miller NA, Farrow EG, Willig LK, Dinwiddie DL, et al. Constellation: a tool for rapid, automated phenotype assignment of a highly polymorphic pharmacogene, CYP2D6, from whole-genome sequences. *NPJ Genom Med*. 2016; 1:15007.
14. Baird JK, Battle KE, Howes RE. Primaquine ineligibility in anti-relapse therapy of *Plasmodium vivax* malaria: the problem of G6PD deficiency and cytochrome P-450 2D6 polymorphisms. *Malar J*. 2018; 17:42.
15. Nelwan EJ, Ekawati LL, Tjahjono B, Setiabudy R, Sutanto I, Chand K, et al. Randomized trial of primaquine hypnozoitocidal efficacy when administered with artemisinin-combined blood schizontocides for radical cure of *Plasmodium vivax* in Indonesia. *BMC Med*. 2015; 13:294.
16. Sutanto I, Tjahjono B, Basri H, Taylor WR, Putri FA, Meilia RA, et al. Randomized, open-label trial of primaquine against vivax malaria relapse in Indonesia. *Antimicrob Agents Chemother*. 2013; 57:1128–35.
17. Baird JK, Louisa M, Noviyanti R, Ekawati L, Elyazar I, Subekti, et al. Association of impaired Cytochrome P450 2D6 activity genotype and phenotype with therapeutic efficacy of primaquine treatment for latent *Plasmodium vivax* JAMA Netw Open. 2018; 1: e181449.
18. Griese EU, Zanger UM, Bruderhans U, Gaedigk A, Mikus G, Morike K, et al. Assessment of the predictive power of genotypes for the in-vivo catalytic function of CYP2D6 in a German population. *Pharmacogenetics*. 1998; 8:15–
19. McElroy S, Sachse C, Brockmoller J, Richmond J, Lira M, Friedman D, et al. CYP2D6 genotyping as an alternative to phenotyping for determination of metabolic status in a clinical trial setting. *AAPS PharmSci*. 2000; 2: E33.
20. Gaedigk A, Bradford LD, Marcucci KA, Leeder JS. Unique CYP2D6 activity distribution and genotype-phenotype discordance in black Americans. *Clin Pharmacol Ther*. 2002; 72:76-89.
21. Lindpaintner K. Pharmacogenetics and the future of medical practice. *Br J Clin Pharmacol*. 2002; 54:221–30.
22. Crews KR, Gaedigk A, Dunnenberger HM, Klein TE, Shen DD, Callaghan JT, et al. Clinical Pharmacogenetics Implementation Consortium (CPIC) guidelines for codeine therapy in the context of cytochrome P450 2D6 (CYP2D6) genotype. *Clin Pharmacol Ther*. 2012; 91:321-326.
23. Crews KR, Gaedigk A, Dunnenberger HM, Leeder JS, Klein TE, Caudle KE, et al. Clinical Pharmacogenetics Implementation Consortium guidelines for cytochrome P450 2D6 genotype and codeine therapy: 2014 update. *Clin Pharmacol Ther*. 2014; 95:376-382.
24. De Gregori M, Allegri M, De Gregori S, Garbin G, Tinelli C, Regazzi M, et al. How and why to screen for CYP2D6 interindividual variability in patients under pharmacological treatments. *Curr Drug Metab*. 2010; 11:276-82.
25. Sistonen J, Fuselli S, Levo A, Sajantila A. CYP2D6 genotyping by a multiplex primer extension reaction. *Clin Chem*. 2005; 51:1291-1295.
26. Sistonen J, Sajantila A, Lao O, Corander J, Barbujani G, Fuselli S. CYP2D6 worldwide genetic variation shows high frequency of altered activity variants and no continental structure. *Pharmacogenet Genomics*. 2007; 17(2):93-
27. Zanger UM, Raimundo S, Eichelbaum M. Cytochrome P450 2D6: overview and update on pharmacology, genetics, biochemistry. *Naunyn Schmiedeberg Arch Pharmacol*. 2004; 369:23-37.
28. Gaedigk A, Simon SD, Pearce RE, Bradford LD, Kennedy MJ, Leeder JS. The CYP2D6 activity score: translating genotype information into a qualitative measure of phenotype. *Clin Pharmacol Ther*. 2008; 83:234-42.

29. Kamenski G, Ayazseven S, Berndt A, Fink W, Kamenski L, Zehetmayer S, et al. Clinical relevance of CYP2D6 polymorphisms in patients of an Austrian Medical Practice: a family practice-based observational study. *Drugs Real World Outcomes*. 2020; 7(1):63-73.
30. Gaedigk A, Sangkuhl K, Whirl-Carrillo M, Klein T, Leeder JS. Prediction of CYP2D6 phenotype from genotype across world populations [published correction appears in *Genet Med*. 2016 Nov;18(11):1167]. *Genet Med*. 2017;19(1):69-76.
31. Caudle KE, Sangkuhl K, Whirl-Carrillo M, Swen JJ, Haidar CE, Klein TE, et al. Standardizing CYP2D6 genotype to phenotype translation: consensus recommendations from the Clinical Pharmacogenetics Implementation Consortium and Dutch Pharmacogenetics Working Group. *Clin Transl Sci*. 2020;13(1):116-124.
32. Swadi AA, Mohammad BI, Hadi NR. Correlation of CYP2D6 allelic polymorphism to outcome of acute coronary syndrome in mid-Euphrates Iraqi patients on metoprolol therapy. *Gene*. 2019; 703:112-119.
33. Gaedigk A, Ingelman-Sundberg M, Miller NA, Leeder JS, Whirl-Carrillo M, Klein TE, et al. The Pharmacogene Variation (PharmVar) Consortium: incorporation of the human Cytochrome P450 (CYP) allele nomenclature database. *Clin Pharmacol Ther*. 2018; 103:399-401.
34. Dong AN, Ahemad N, Pan Y, Palanisamy UD, Yiap BC, Ong CE. Functional and structural characterisation of common cytochrome P450 2D6 allelic variants-roles of Pro34 and Thr107 in catalysis and inhibition. *Naunyn Schmiedebergs Arch Pharmacol*. 2019; 392:1015-1029.
35. St Jean PL, Xue Z, Carter N, Koh GC, Duparc S, Taylor M, et al. Tafenoquine treatment of *Plasmodium vivax* malaria: suggestive evidence that CYP2D6 reduced metabolism is not associated with relapse in the Phase 2b DETECTIVE trial. *Malar J*. 2016; 15:97.
36. Brasil LW, Rodrigues-Soares F, Santoro AB, Almeida ACG, Kuhn A, Ramasawmy R, et al. CYP2D6 activity and the risk of recurrence of *Plasmodium vivax* malaria in the Brazilian Amazon: a prospective cohort study. *Malar J*. 2018; 17:57.
37. Silvino AC, Costa GL, Araujo FC, Ascher DB, Pires DE, Fontes CJ, et al. Variation in human cytochrome P-450 drug-metabolism genes: a gateway to the understanding of *Plasmodium vivax* PLoS ONE. 2016; 11: e0160172.
38. Raimundo S, Toscano C, Klein K, Fischer J, Griese EU, Eichelbaum M, et al. A novel intronic mutation, 2988GA, with high predictivity for impaired function of cytochrome P450 2D6 in white subjects. *Clin Pharmacol Ther*. 2004; 76:128-38.
39. Wang D, Poi MJ, Sun X, Gaedigk A, Leeder JS, Sadee W. Common CYP2D6 polymorphisms affecting alternative splicing and transcription: long-range haplotypes with two regulatory variants modulate CYP2D6 activity. *Hum Mol Genet*. 2014; 23:268-278.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Figures



Figure 1

1 The map of China highlighting the geographic location of Yunnan Province. Different colors represent different administrative areas, and green labels indicate the main source of infection in areas with many imported cases. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

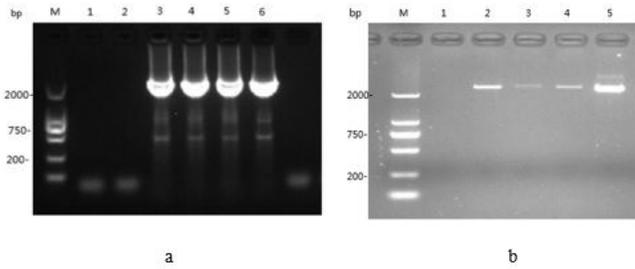


Figure 2
 PCR products of CYP2D6 gene fragment. a M: DNA marker; 1: blank control of PCR for the first round; 2: blank control of PCR for the second round; 3-6: The amplification products of exon 1-4 in CYP2D6 gene; b M: DNA marker; 1: blank control of PCR; 2-5: The amplification products of exon 5-9 in CYP2D6 gene;

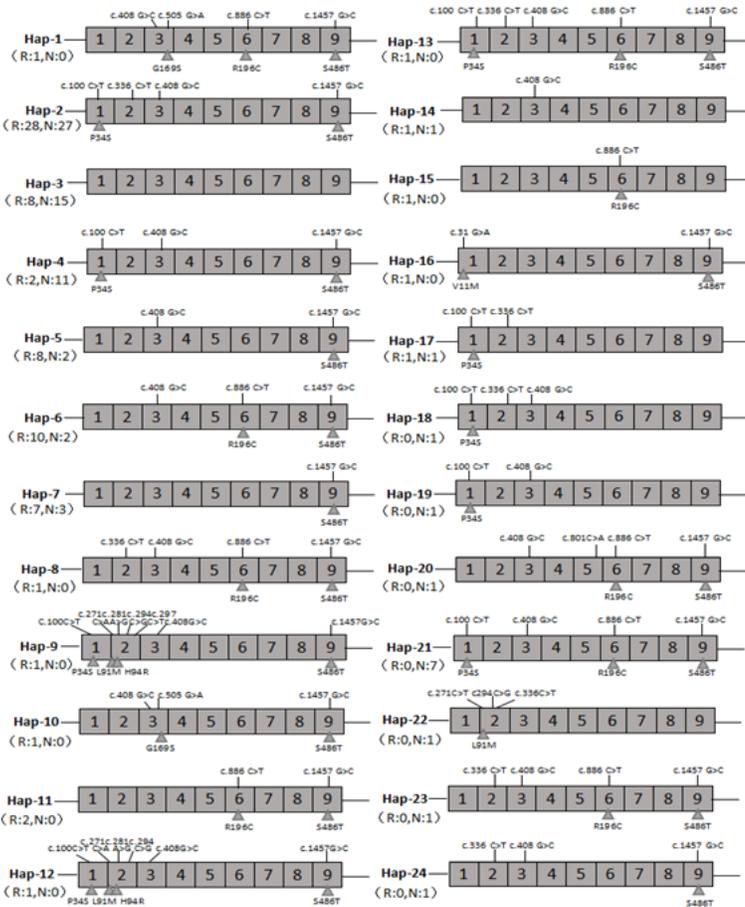


Figure 3
 The 9 exons are indicated by numbered boxes with DNA polymorphisms indicated on top. Predicted amino acid changes are indicated below; no change is synonymous mutation. R: Relapsed cases; N: Non-relapsed cases.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Figure1.jpg](#)