Genetic polymorphism analysis of cytochrome P4502E1 (CYP2E1) in a Chinese Tibetan population

Li Wang, MD<sup>a,b,c</sup>, Guoxia Ren, MD<sup>d,e</sup>, Jingjie Li, MS<sup>f</sup>, Linhao Zhu, MM<sup>a,b,c</sup>, Fanglin Niu, MS<sup>f</sup>, Mengdan Yan, MS<sup>i</sup>, Jing Li, MS<sup>i</sup>, Dongya Yuan, MD<sup>a,b,c</sup>,*, Tianbo Jin, MD<sup>a,b,c,f</sup,*

Abstract

Cytochrome P4502E1 (CYP2E1) gene polymorphisms vary markedly in frequency among different ethnic and racial groups. We studied the genotype distributions and allele frequencies of 3 CYP2E1 polymorphisms: CYP2E1<sup>*1A</sup>, CYP2E1<sup>*7A</sup>, and CYP2E1<sup>*7C</sup> by polymerase chain reaction technique in a sample of 100 healthy subjects representing Tibetan population.

The frequencies of CYP2E1<sup>*1A</sup>, *7A, and *7C alleles were 0.705, 0.125, and 0.170, respectively. Compared with other populations, we found that the allele frequencies of the variants −352A>G (rs2070672) and −333A>T (rs2070673) in this Tibetan population have significant differences compared with European-American, African-American, Japanese, Korean, and other different geographic areas in Chinese Han population. Furthermore, the results of protein prediction revealed that the variant 6397G>A (rs61710826) could influence the protein structure and function.

These findings in this study would be valuable for pharmacogenetics for drug therapy and drug discovery. However, further studies in larger samples are warranted to confirm our results.

Abbreviations: CYP2E1 = cytochrome P4502E1, DNA = deoxyribonucleic acid, EDTA = ethylene diamine tetraacetic acid, PCR = polymerase chain reaction, PolyPhen-2 = polymorphism phenotyping v2, SIFT = sorting intolerant from tolerant.

Keywords: CYP2E1 gene, genetic polymorphism, Tibetan population

1. Introduction

The cytochrome P450 (CYP450) enzymes play a central role in the oxidative metabolisms of exogenous and endogenous compounds, including drugs, food additives, industrial solvents, and pollutants, converting them to reactive metabolites.[1] Besides detoxification, many CYP isoforms catalyze the metabolic activation of procarcinogens to their ultimate carcinogenic forms.[2] However, CYP450 enzymes show extensive structural differences due to genetic polymorphisms in the corresponding genes, which give rise to the absence of gene product, enzymes with increased, reduced or altered activity, or alteration in enzyme regulation, which may be responsible for interindividual and interethnic variabilities in drug response and carcinogenetic susceptibility.[3–5] Therefore, the individualized drug therapy based on genotype analysis can effectively reduce adverse effects and improve drug efficacy. Some genotype–phenotype analysis of CYP450 gene, such as CYP3A4,[6] and CYP2C19[7,8] have been reported in Chinese minority population.

The cytochrome P4502E1 (CYP2E1), as a member of the cytochrome P450 superfamily, is responsible for the metabolic activation of many low-molecular weight compounds, such as ethanol, benzene, vinyl chloride, and N-nitrosamines.[9,10] Compared with other cytochromes P450, the CYP2E1 enzyme has a relatively high redox potential and can induce peroxidation (lipid and NADPH dependent) or other oxidative stress causing the production of reactive oxygen species.[11,12] Therefore, the CYP2E1 enzyme is considered an important source of reactive oxygen species in alcohol-induced liver injury.[13] In addition, it has been reported that endogenous and exogenous substrates, which might be associated with human susceptibility to toxicity and carcinogenicity caused by industrial and environmental chemicals, can regulate the CYP2E1 level.[14] The human CYP2E1 gene is located in 10q24.3–qter region of chromosome 10.[15] Some polymorphisms in the CYP2E1 have been reported to be associated with the risk of cancer[16,17] and other diseases.[18,19] However, previous studies demonstrated that the alleles and genotypes frequencies of CYP2E1 polymorphisms...
have significant differences among different ethnic and racial groups.\(^{[20-22]}\) The CYP2E1 polymorphisms can cause the differences of interindividual drug metabolism and liver injury, or even severe adverse drug reaction.

The Tibetan population is a minority ethnic group in China with unique lifestyle, diverse genetic background, dietary habit, culture, and geographical environment. According to 2014 Census, Tibetans with a population of 2.2 million live mostly in the Tibet Autonomous Region and the 10 Tibetan Autonomous Prefectures in Gansu Qinghai, Sichuan, and Yunnan provinces of China. We systematically screened some CYP2E1 gene polymorphisms of 100 healthy, unrelated Tibetans for polymorphisms and compared the allelic frequencies with those in the ethnic China population. Our study hope to find corresponding phenotypes and offer recommendations pertaining to the drug substrates of CYP2E1 in the Tibetan population.

2. Materials and methods

2.1. Subjects

A total of 100 unrelated Tibetan Chinese healthy subjects consisted of 50 males and 50 females were recruited for the population genetics study. The subjects selected were deemed healthy based on their medical history and a physical examination. All participants resided in the Xinjiang Autonomous Region of China and had the Tibetan paternal ancestry for at least 3 generations. The study protocol was approved by the Ethics Committee of Xizang Minzu University and performed in accordance with the Declaration of Helsinki. We informed each subject about the experimental procedures and the purpose of the study and written informed consent was obtained from all participants before enrollment in the study.

2.2. PCR and DNA sequencing

Five milliliters peripheral venous blood sample was collected from each participant in EDTA (ethylene diamine tetraacetic acid) containing vacuumed tube, and the genomic DNA (deoxyribonucleic acid) was extracted from the of the using the GoldMag-Mini Whole Blood Genomic DNA Purification Kit (GoldMag Ltd., Xi’an, China) according to the manufacturer’s instructions. The primers for polymerase chain reaction (PCR) were designed to amplify the exons and the 3’-untranslated region of the CYP2E1 gene using a standard procedure, and the sequences are listed in Table 1.

PCR was performed in a total volume of 10\(\mu\)L reactions containing 3\(\mu\)L HotStar Taq Master Mix, 1\(\mu\)L of template DNA, 0.5\(\mu\)L each primer (5\(\mu\)M), and 3\(\mu\)L deionized water. Thermal cycling conditions were at 95°C for 15 minutes (denaturation), followed by 35 cycles at 95°C for 30 seconds (denaturation), 55°C to 64°C for 30 seconds (annealing), 72°C for 1 minute (extension), and a final extension at 72°C for 3 minutes to hold. The PCR products were purified by incubating with 0.5\(\mu\)l shrimp alkaline phosphate (Roche, Basel, Switzerland), 1.5\(\mu\)L deionized water, and 8\(\mu\)L HotStar PCR product, for a total volume of 10\(\mu\)L, at 38°C for 30 minutes, followed by heat inactivation at 80°C for 15 minutes. The purified PCR products were directly sequenced using an ABI Prism BigDye Terminator Cycle Sequencing Kit version 3.1 (Applied Biosystems Inc., Foster City, CA) on an ABI Prism3100 sequencer (Applied Biosystems). We used specific primers to detect deletions and duplications of CYP2E1.

2.3. Data analysis

The initial analyses of the sequences including base calling, fragment assembly, and detection of SNPs, insertions, and deletions were used in the Sequencer 4.10.1 (http://www.genecodes.com/) software. The Human Cytochrome P450 (CYP) Allele Nomenclature Database describes CYP2E1 variants according to the NCBI reference sequence: NC_000010.11 and CYP allele nomenclature (http://www.cypalleles.ki.se/). The differences of allelic frequencies between Tibetan and other ethnic populations were compared using the \(\chi^2\) test with a significance level set at 0.05. Hardy–Weinberg equilibrium for each genetic variant and Linkage disequilibrium (LD) between loci pairs were assessed using Haploview software (version 4.2).\(^{[23]}\) The haplotypes were constructed from the selected SNPs and the haplotype frequencies were derived for the Tibetan population.

2.4. Transcriptional prediction

We predicted the protein function of nonsynonymous SNPs (nsSNPs) in CYP2E1 coding regions used the online tools SIFT (Sorting Intolerant from Tolerant, http://sift.jcvi.org/) and PolyPhen-2 (Polymorphism Phenotyping v2, http://genetics.bwh.harvard.edu/pph2/). The SIFT results were divided into 4 categories based on these scores: tolerant (0.00–0.10), border-line (0.10–0.20), potentially intolerant (0.051–0.10), and intolerant (0.050–0.05).\(^{[24]}\) PolyPhen-2 results were divided into 5 categories: benign (0.00–0.15), possibly damaging (0.15–0.85), and probably damaging (0.85–1.00).\(^{[25]}\) The prediction accuracy of SIFT and PolyPhen-2 was 63% and 75%, the false positive rate is 19% and 9%, respectively.

3. Results

3.1. Genetic variants

We sequenced CYP2E1 from our study subjects and successfully identified a total of 23 CYP2E1 polymorphisms in the Tibetan population. Three polymorphisms were not previously reported in either the NCBI database or the Human Cytochrome P450 430 Allele Nomenclature Committee tables. Two of the novel polymorphisms (971C>T and 6179C>T) were within the introns, and the other (11398C>A) was in the 3’UTR, as shown in Table 2.

3.2. Alleles and genotypes frequencies

Three alleles and 3 genotypes of CYP2E1 were identified on the basis of the polymorphisms found in the Chinese Tibetan
population (Table 3). All the CYP2E1 alleles and genotype frequencies were in the Hardy–Weinberg equilibrium. The most frequent alleles in the Tibetan populations were the wild-type allele CYP2E1*1A (70.5%), followed by the CYP2E1*7C allele (17.0%) and the CYP2E1*7A allele (12.5%). Individuals with the wild-type *1A/*1A genotype have normal enzyme activity, and this genotype was the most prevalent (41.0%) in the Chinese Tibetan population. Other identified genotypes included the heterozygous *1A/*7C genotype (34.0%) and the *1A/*7A genotype (25.0%).

### Table 2

| SNP       | New alleles | Gene position | Region       | Nucleotide change | Amino-acid effect | Frequency (%) | Flanking sequence |
|-----------|-------------|---------------|--------------|-------------------|-------------------|---------------|-------------------|
| rs2070672 | *7C         | −352 Promoter | A>G          | No translated     | 35                | CCGTGGTCTCATGACGGA |                  |
| rs2070673 | *7A, 7C     | −333 Promoter | A>T          | No translated     | 59                | AAAAGGCAGSGWGGCTAGCCCATCA |                  |
| rs1499408 | Novel       | 971 Intron 1  | C>A          | No translated     | 4                 | GTTCCCTCCATGCTCGTG |                  |
| rs2070674 |            | 10463 Exon 2  | T>C          | No translated     | 2                 | TCCGGTCCTCTACCACGGA |                  |
| rs2070675 |            | 10275 Intron 7 | A>G          | No translated     | 31                | ATTATAGCAATGCCCACGAGGA |                  |
| rs2070676 |            | 10238 Intron 7 | C>T          | No translated     | 3                 | TCCGTTCCTGCAATGCTCAG |                  |
| rs2070677 |            | 9926 Intron 7  | C>T          | No translated     | 1                 | AACATTGGTTCAGGGGTTT |                  |
| rs2070678 |            | 6634 Intron 6  | A>G          | No translated     | 1                 | GTAGGACGAGGGAAGGGGA |                  |
| rs2837174 |            | 6444 Exon 6   | A>G          | No translated     | 100               | TCGAGCCGAGGAGGAGGAGGAGGA |                  |
| rs2480257 |            | 11610 3'UTR    | G>T          | No translated     | 3                 | TTCTGGGGGTACGCTTCAC |                  |
| rs2515641 |            | 4441 Intron 2  | C>T          | No translated     | 31                | TCCCTCCAAGCCCTCTACCAA |                  |
| rs4129940 |            | 9972 Intron 2  | C>T          | No translated     | 3                 | TCTTTTCAAGAAGCAAGAGGA |                  |

### Table 3

| Allele | Total (n=200) | Phenotype | Frequency (%) |
|--------|---------------|-----------|---------------|
| *1A    | 141           | Normal    | 70.5          |
| *7A    | 25            | Normal    | 12.5          |
| *7C    | 34            | Normal    | 17.0          |

| Genotype | Total (n=100) | Phenotype | Frequency (%) |
|----------|---------------|-----------|---------------|
| *1A/*1A  | 41            | Normal    | 41.0          |
| *1A/*7A  | 25            | Normal    | 25.0          |
| *1A/*7C  | 34            | Normal    | 34.0          |

3.4. Interpopulation comparisons

We further compared CYP2E1 polymorphisms distribution patterns between Tibetan population and previously published data from different countries in China. The results showed that the frequencies of the −352A>G (CYP2E1*7C) and the −333A>T (CYP2E1*7A and *7C) in our study group were significantly higher (P < 0.05) than most of other groups. Furthermore, compared them with the frequencies in other ethnic populations, we found that the −352A>G and the −333A>T were significant differences compared with African-American, European-American, and Korean, as shown in Table 4.

3.5. Predicted protein function of the nonsynonymous mutation

The protein prediction results of the nonsynonymous variant 6397G>A from the SIFT analysis indicated that substitution at position 288 from Gly to Ser is predicted to be tolerated with a score of 0.25. The result performed by PolyPhen-2 analysis showed that this mutation is predicted to be probably damaging with a score of 0.945 (HumDiv) and a score of 0.613 (HumVar), as shown Figure 2.

4. Discussion

The cytochrome P450 isozyme CYP2E1 are highly relevant in the metabolism of many low-molecular weight drugs, toxicants, and carcinogens, and may give rise to important interindividual and interethnic differences in patient responsiveness and adverse drug reactions. CYP2E1 gene polymorphisms could be associated with
the high degree of individual variability in the susceptibility to developing cancer and other diseases related, since polymorphic alleles encode proteins with altered catalytic activities or show differences in gene expression. To better understand the distribution of CYP2E1 allele and genotype frequencies in the Tibetan populations, we systematically screened the whole CYP2E1 polymorphisms from 100 healthy, unrelated Tibetans. We identified 22 genetic variants including 3 novel polymorphisms, 3 alleles (∗1A, ∗7A, and ∗7C), and 3 genotypes (∗1A/∗1A, ∗1A/∗7A, and ∗1A/∗7C) of CYP2E1 in our study Tibetan Chinese population. We also compared 2 major allelic polymorphisms (−352A>G, SNP rs2070672 and −333A>T, SNP rs2070673) with previous observations of other ethnic populations and found that −352A>G and −333A>T were different from those of European-American, African-American, Japanese, Korean, and other different geographic areas of Mainland China Han populations. These differences of polymorphisms distribution could be attributed to the origin and geographical isolation experienced by different ethnic populations, as well as their dietary habits and lifestyles, all of which may affect CYP2E1 polymorphisms. However, no significant difference was found between −352A>G and −352A>G in the Tibetan and Han Chinese populations. It is probably because that the sample size is relatively small for a population genetics research and the reference studies have variable quality. In addition, it has been reported that −333A>T exists at similar frequencies in European population, with the highest frequency among Africans.

The protein prediction results revealed that the variant 6397G>A could influence the protein structure and function, and the results of SIFT (tolerant, score = 0.25) and PolyPhen-2 (damaging, score = 0.945) were inconsistent. The inconsistency of functional prediction may be due to that different algorithms are based on different training data, each has its own strength and weakness. In addition, a bioinformatics study of phenotype prediction of deleterious nonsynonymous single-nucleotide polymorphisms in human alcohol metabolism-related genes showed that the mutation 6397G>A is predicted to be borderline (SIFT, score = 0.18) and benign (PolyPhen-2, score = 0.254). The reason for the difference is not clear. The results identified here should be confirmed by other means in further studies.

**Figure 1.** Linkage disequilibrium analysis of CYP2E1. LD is displayed by standard color schemes, with bright red for very strong, and blue for intermediate LD, and white for no LD.

**Table 4**

| Population       | Total number | Allele frequencies (%) | References |
|------------------|--------------|------------------------|------------|
| Tibetan          | 100          | 35.0                   | 59.0       | [26]       |
| Shantou          | 100          | 18.8                   | 45.3       |
| Shanghai         | 100          | 14.1                   | 40.1       |
| Shenyang         | 100          | 18.8                   | 35.9       |
| Xi’an            | 100          | 21.0                   | 36.5       |
| Chinese          | 400          | 18.4                   | 46.7       | [21]       |
| European         | 268          | —                      | 17.0       |
| African          | 268          | —                      | 80.0       |
| African-American | 48           | 11.5                   | 26.0       | [24]       |
| European-American| 48           | 2.1                    | 15.6       |
| Han Chinese      | 48           | 25.0                   | 46.9       |
| Japanese         | 48           | 22.9                   | 44.8       |
| Korean           | 96           | 21.8                   | 40.4       |

* P < 0.05, compared with the data of the present study.
** P < 0.01, compared with the data of the present study.
5. Conclusions

In conclusion, our results provide new information regarding CYP2E1 genetic polymorphisms in Chinese Tibetan population. These findings in this study would be value for further study in assessing the susceptibility of different populations to diseases related to CYP2E1 polymorphisms and determining whether it is necessary to design different therapeutics and toxicological protocols to reduce the risk of population, including pharmacogenetics for drug therapy and drug discovery. Further studies in larger groups are warranted to confirm our results.

Acknowledgments

We are grateful to all the participants in this study.

References

[1] Nebert DW, Russell DW. Clinical importance of the cytochromes P450. Lancet 2002;360:1135–62.
[2] Guengerich FP, Shimada T. Oxidation of toxic and carcinogenic chemicals by human cytochrome P-450 enzymes. Chem Res Toxicol 1991;4:391–407.
[3] Daly AK. Pharmacogenetics of the cytochromes P450. Curr Top Med Chem 2004;4:1733–44.
[4] Vines P. The relationship between polymorphisms of xenobiotic metabolizing enzymes and susceptibility to cancer. Toxicology 2002; 181–182:457–62.
[5] Evans WE. Pharmacogenomics: marshalling the human genome to individualise drug therapy. Gut 2003;52(Suppl 2):ii10–8.
[6] Jin T, Yang H, Zhang J, et al. Polymorphisms and phenotypic analysis of cytochrome P450 3A4 in the Uygur population in northwest China. Int J Clin Exp Pathol 2015;8:7083.
[7] Jin T, Zhang X, Meng T, et al. Genotype-phenotype analysis of CYP2C19 in the Tibetan population and its potential clinical implications in drug therapy. Mol Med Rep 2016;13:2117–23.
[8] Ding Y, Xu D, Zhang X, et al. Genetic polymorphisms and phenotypic analysis of drug-metabolizing enzyme CYP2C19 in a Li Chinese population. Int J Clin Exp Pathol 2015;8:13201.
[9] Guengerich FP, Kim DH, Iwasaki M. Role of human cytochrome P-450 IIE1 in the oxidation of many low molecular weight cancer suscept. Chem Res Toxicol 1991;4:168–79.
[10] Arinc E, Arslan S, Bozcaarmutlu A, et al. Effects of diabetes on rabbit kidney and lung CYP2E1 and CYP2B4 expression and drug metabolism and potentiation of carcinogenic activity of N-nitrosodimethylamine in kidney and lung. Food Chem 2007;45:107–18.
[11] Ingelman-Sundberg M, Johansson I, Yin H, et al. Ethanol-inducible cytochrome P4502E1: genetic polymorphism, regulation, and possible role in the etiology of alcohol-induced liver disease. Alcohol 1993;10:447–52.
[12] Zhukov A, Ingelman-Sundberg M. Relationship between cytochrome P450 catalytic cycling and stability: fast degradation of N-heptanoyl-inducible cytochrome P450 2E1 (CYP2E1) in hepatoma cells is abolished by inactivation of its electron donor NADPH-cytochrome P450 reductase. Biochem J 1999;340(Pt. 2):453.
[13] Cederbaum AI. CYP2E1–biological, genotoxic, and toxicological aspects. J Toxicol Environ Health 2003;76:174–85.
[14] Umeno M, Mcbride OW, Yang CS, et al. Human ethanol-inducible P450IIIE1: complete gene sequence, promoter characterization, chromosome mapping, and cDNA-directed expression. Biochemistry 1988;27:9006–13.
[15] Zhang Y, Hughes KJ, Zahm SH, et al. Genetic variations in xenobiotic metabolic pathway genes, personal hair dye use, and risk of non-Hodgkin lymphoma. Am J Epidemiol 2009;170:1222–30.
[16] Barry KH, Zhang Y, Lan Q, et al. Genetic variation in metabolic genes, occupational solvent exposure, and risk of non-Hodgkin lymphoma. Am J Epidemiol 2011;173:404–13.
[17] Huo R, Tang K, Wei Z, et al. Genetic polymorphisms in CYP2E1: association with schizophrenia susceptibility and risperidone response in the Chinese Han population. PLoS One 2012;7:e34809.
[18] Liao LH, Zhang H, Lu MP, et al. Single-nucleotide polymorphisms and haplotype of CYP2E1 gene associated with systemic lupus erythematosus in Chinese population. Arthritis Res Ther 2011;13:R11.
[19] Huang X, Chen L, Song W, et al. Systematic functional characterization of cytochrome P450 2E1 promoter variants in the Chinese Han population. PLoS One 2012;7:e40883.
[20] Bonifazpeña V, Contreras AV, Struchiner CJ, et al. Exploring the distribution of genetic markers of pharmacogenomics relevance in Brazilian and Mexican populations. PLoS One 2014;9:e112640.
[21] Kim JH, Chrong HS, Park BL, et al. Direct sequencing and comprehensive screening of genetic polymorphisms on CYP family genes (CYP2A6, CYP2B6, CYP2C8, and CYP2E1) in five ethnic populations. Arch Pharm Res 2015;38:115–28.
[22] Barret JC, Trey B, Muller J, et al. Haplomview: analysis and visualization of LD and haplotype maps. Bioinformatics 2005;21:1263–5.
[23] Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc 2009;4:1073–81.
[24] Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. Nat Methods 2010;7:304–9.
[25] Tang K, Li X, Xing Q, et al. Genetic polymorphism analysis of cytochrome P4502E1 (CYP2E1) in Chinese Han populations from four different geographic areas of Mainland China. Genomics 2010;95:224–9.
[26] Wang LL, Yang AK, Li Y, et al. Phenotype prediction of deleterious non-synonymous single nucleotide polymorphisms in human alcohol metabolism-related genes: a bioinformatics study. Alcohol 2010;44:425–38.