0031-3998/10/6703-0280
PEDIATRIC RESEARCH

Vol. 67, No. 3, 2010
Printed in U.S.A.

Mutational Spectrum of Phenylketonuria in the Chinese Han Population: A Novel Insight into the Geographic Distribution of the Common Mutations

TIANWEN ZHU, SHENGYING QIN, JUN YE, WENJUAN QIU, LIANSHU HAN, YAFEN ZHANG, AND XUEFAN GU

Department of Endocrinology and Genetic Metabolism [T.Z., J.Y., W.Q., L.H., Y.Z., X.G.] and Shanghai Institute of Pediatric Research [J.Y., W.Q., L.H., Y.Z., X.G.], Shanghai Jiao Tong University School of Medicine, Shanghai 200092, China; Bio-X life Science Research Center [S.Q.], Shanghai Jiao Tong University, Shanghai 200030, China

ABSTRACT: The absence of a comprehensive analysis for phenylketonuria (PKU) mutations in the Chinese Han population has resulted in continued studies during the past 18 y to elucidate the mutational spectrum in patients from virtually all Chinese regions. Our study systematically investigated 13 exons and their surrounding introns of the phenylalanine hydroxylase (*PAH*) gene in 212 unrelated patients using PCR and direct sequencing. A total of 79 different mutations were identified in 405 of 424 mutant PAH alleles including 15 novel ones. Eight mutations, R243Q, Ex6–96A>G, IVS4 − 1G>A, R413P, Y356X, R111X, R241C, and V399V, with a relative frequency of 3% or more, accounted for two thirds of the identified ones. The data presented in this study indicates that the total pool of mutant PAH alleles in China consisted of a small number of common mutations and a very high number of rare mutations. Moreover, by merging the findings of previous studies to generate a more composite data set for the Chinese mainland, it is shown that there are no significant differences of the common mutations between southern and northern except for R413P statistically, raising questions about the previous hypothesis that great variations on mutation frequencies exist between above regions. **(*Pediatr Res* 67: 280–285, 2010)**

Phenylketonuria (PKU; MIM# 261600) is a common autosomal recessive inborn error of amino acid metabolism and mainly results from mutations of the phenylalanine hydroxylase gene (*PAH*; 612349). So far, >500 different mutant alleles have been identified at the *PAH* locus and listed in the *PAH* Mutation Analysis Consortium database (http://www.pahdb.mcgill.ca/), which cause different levels of reduction in the catalytic activity of the enzyme, generating a wide spectrum of biochemical and clinical phenotypes (1,2). Data (3–18) on the distribution and relative frequencies of the mutations have been described for various populations and have shown great variability in the mutational spectrum and differences in the degree of heterogeneity, which are useful for further understanding of both the clinical features and the population genetics of the disorder.

To date, there has been no comprehensive population genetic study of PKU focused on the Han Chinese. This largest ethnic population in China is naturally divided into two groups, the Southern and Northern Han, by the Yangtze River (19,20), resulting in the formation of different founder populations with relatively isolated consanguinity. It is believed that the difference between two groups is greater than that between given subpopulations and ethnic minorities at the same location, and the stratification would affect the mutational spectrum. Based on previous studies (21–23), there are some indicators to suggest that regional variations on mutation frequencies do exist between southern and northern China. However, the evidence to support this statement is limited, as existing ones are selective and of small sample size (7,24–30).

We undertook this study with the objectives of reaching full ascertainment of PKU mutations in the Chinese Han and of investigating regional differences in the mutation spectrum within China, looking in greater depth at the possible explanations for the geographic distribution of the common mutations.

MATERIALS AND METHODS

Subjects. A total of 212 unrelated Chinese Han patients with PKU, corresponding to 424 independent alleles, were investigated. The geographical distribution of mutant alleles within China was defined on the basis of the origin of the birth parents of each case studied. So, only those patients with both parents originating from the same native place were recruited. Therefore, 112 came from southern China and 100 from northern.

Most (150) of the patients, accounting for 71%, were identified when they showed mental retardation between 6 mo and 3 y olds; the remaining (62), accounting for 29%, in neonatal screening. Their phenotypes were classified based on the pretreatment plasma phenylalanine (phe) levels or the phe level at diagnosis. Accordingly, 113 patients were classic PKU (phe >1200 μM/L) (113/212); 66, mild PKU (phe 600~1200 μM/L) (66/212); 15, MHP (phe 120~600 μM/L) (15/212); 18 cases were unclassified, with their phe level unavailable. An explanation of the study was given to the participating patients and a standard informed consent, which was reviewed and approved by the Shanghai Ethical Committee of Human Genetic Resources, was obtained from all subjects.

Polymerase chain reaction condition and DNA sequencing. Systematic mutation screening was performed by direct sequencing. Genomic DNA was isolated from the peripheral blood using the standard procedure (31). PCR primers were designed to amplify all 13 exons and surrounding introns of the PAH gene. The primers used in this study are shown in Table 1. The PCRs

Received June 16, 2009; accepted October 19, 2009.
Correspondence: Xuefan Gu, M.D., Ph.D., Department of Endocrinology and Genetic Metabolism, Xin-Hua Hospital, Shanghai Institute for Pediatric Research, No 1665 Kongjiang Road, Shanghai 200092, China; e-mail: guxuefan@online.sh.cn
T.Z. and S.Q. contributed equally to this work.
Supported by Hi-Tech Research and Development Program (2007AA02Z447), National Science and Technology Program (2006BAI05A05 and 2006BAI05A07), and Shanghai Bureau of health (2008ZD001).

Abbreviations: PAH, phenylalanine hydroxylase; **phe,** phenylalanine; **PKU,** phenylketonuria

Table 1. *PAH primers and amplifications*

Region	Primer	Primer sequence	PCR product size (bp)
Exon 1	EXF1	5′-CTCCCTAGTGCGAGGTTAA-3′	246
	EXR1	5′-CAGGAAGCACCAGCAGTC-3′	
Exon 2	EXF2	5′-ATTTAATTGCCCTGGACTTACT-3′	411
	EXR2	5′-TCAAATTCAAATCTGCCTGTTC-3′	
Exon 3	EXF3	5′-GCGTTAGGTTTTCCTGTTCT-3′	472
	EXR3	5′-TATGTCCACTCATTTAATCCC-3′	
Exon 4	EXF4	5′-AGAGAACATCTCCAGGAAAGA-3′	456
	EXR4	5′-CCCAGCCCTCGTGTAAATAG-3′	
Exon 5	EXF5	5′-AGGGTATAACCAAGGGAAGG-3′	264
	EXR5	5′-ATGAGGGCAAGGGAGAAGC-3′	
Exon 6	EXF6	5′-CCCCGACTCCCTCTGCTAA-3′	339
	EXR6	5′-CCTCTGCCTCAATCCTCCC-3′	
Exon 7	EXF7	5′-CTGCCTAGCGTCAAAGCC-3′	351
	EXR7	5′-CAGCCAGCAATGAACCCAAA-3′	
Exon 8	EXF8	5′-ACTGAGTCTGGCTTGGCTTAA-3′	243
	EXR8	5′-ACCTGGTTTCCGCTCTTGC-3′	
Exon 9	EXF9	5′-GATACTAACCGACCCTGTG-3′	389
	EXR9	5′-ATAGCACTCCACCATCCAC-3′	
Exon 10	EXF10	5′-AGGTATCCCTTCATCCAGTC-3′	348
	EXR10	5′-CCCACAGCCATCATCAAATC-3′	
Exon 11	EXF11	5′-ATTTGGGCTGTGATGTAGAAG-3′	510
	EXR11	5′-TTCAGTGTCTTGACTTGGTGG-3′	
Exon 12	EXF12	5′-CTAGGGAGGTGTCCGTGTT-3′	299
	EXR12	5′-GGCGATGGTAGGGAAAGAC-3′	
Exon 13	EXF13	5′-TCCAAGAAGCCCACTTATCC-3′	375
	EXR13	5′-TCTCTAAATCAAAGATGACCC-3′	

PAH, phenylalanine hydroxylase; F, forward; R, reverse.

were carried out on the Gene Amp PCR system 9700 (Applied Biosystems, CA), with a cycling protocol which consisted of denaturation at 95°C for 30 s, 50–65°C for 1 min, and 72°C for 10 min. Amplified DNAs were incubated with shrimp alkaline phosphatase (Roche, Basel, Switzerland) and exonulease (New England Biolabs Inc., MA) at 37°C for 45 min. The products were sequenced using an ABI Prism BigDye Terminator Cycle Sequencing Kit, version 3.1 (Applied Biosystems) on an ABI Prism 3100 sequencer.

Nomenclature. It has been the convention in the *PAH* Mutation Analysis Consortium to use "trivial names" (32,33). The corresponding systematic names are given in the database as well (http://www.pahdb.mcgill.ca/). These two types of nomenclature have been used throughout our study.

Calculation of homozygosity. Homozygosity (j) at the *PAH* locus in the population was determined by $j = \sum x_i^2$, where x_i was the frequency of the ith allele. Here each of the uncharacterized alleles was defined as having a frequency of 1/N, where *N* was the total number of mutant chromosomes investigated (15).

Statistical analysis. Mutation and genotype frequencies were calculated by the counting method. Comparisons of them between two geographic populations were done using χ^2 tests or Fisher's exact tests with a significance level set at 0.05. All statistical calculations were computed with scripts running on an SPSS 13.0 platform.

Automated splice site analysis. Reference and variant genomic sequences were used to predict splice sites to evaluate potential splice site variants (http://www.fruitfly.org/seq_tools/splice.html).

RESULTS

Mutational spectrum. Mutation analysis was performed on 212 PKU individuals, representing 424 independent mutant chromosomes, and a potential disease-causing mutation was identified on 405 of 424, corresponding to a mutation detection rate of 95%. The spectrum was composed of 79 different mutations. The majority were missense mutations (54 of 79, 68.4%), with 12 splice-site ones, 10 nonsense ones, and three deletion ones, the latter three types classified as null mutation. These mutations were distributed across the *PAH* coding sequence except exons 1 and 13, with 30% (24 of 79) occurring in exon 7. A substantial proportion mutant alleles (62%) were account for by R243Q (26%), Ex6–96A>G (9%), IVS4 − 1G>A (6%), R413P (5%), Y356X (5%), R111X (3.7%), R241C (3.7%), and V399V (3.2%). All other mutant alleles were present at relative frequencies of 2.5% or less. The relative frequencies of mutations found in our study are summarized in Table 2. The geographic distribution of mutant alleles on Chinese mainland was also investigated, and we observed that a uniform distribution of R243Q, IVS4 − 1G>A, Y356X, R241C, V399V, Ex6–96A>G, and R111X, the prevalent mutations in China, with only one common mutation, R413P, although present overall, clustering in northern China.

Novel mutations. We found 15 novel mutations, F121L, Y154C, A156P, E183G, L227Q, E228X, R270G, P275A, I324N, C357Y, C357X, P362L, Y414X, IVS4 + 2T>A, IVS5 − 2A>G, and two novel polymorphisms, IVS6 − 59C>G and IVS3 + 164T>A, not recorded in the PAH mutation Analysis Consortium database. Four novel sequence variations were detected in intron; the other 13 were in coding region. Reference and variant genomic sequences were used to predict splice sites to evaluate potential splice site variants (http://www.fruitfly.org/seq_tools/splice.html), and IVS4 + 2T>A, IVS5 − 2A>G were found to alter the existing splice sites. Those novel single nucleotide changes, F121L, Y154C, A156P, E183G, L227Q, E228X, R270G, P275A, I324N, C357Y, C357X, P362L, and Y414X, which occurred in coding region, altered the original amino acid sequence (34), and we assumed that these variants are functionally relevant although this needs to be confirmed by additional studies. Except for IVS6 − 59C>G and IVS3 + 164T>A, two novel

Table 2. *Spectrum of PAH mutations detected in Chinese Han population*

Trivial name (protein effect)	Systematic name (DNA level)	Exon	Characters of mutation	Relative frequency(%)			
				Nothern China (n = 100)	Southern China (n = 112)	Chinese (n = 212)	p*
R243Q	c.728G>A	7	Missense	0.2550	0.2321	0.2543	0.6616
EX6–96A>G	c.611A>G	6	Splice	0.0650	0.1027	0.0889	0.1463
IVS4–1G>A	c.442–1G>A	Intron 4	Splice	0.0600	0.0625	0.0642	0.8742
R413P	c.1238G>C	12	Missense	0.0800	0.0268	0.0543	0.0155
Y356X	c.1068C>A	11	Nonsense	0.0550	0.0446	0.0519	0.6560
R111X	c.331C>T	3	Nonsense	0.0400	0.0313	0.0370	0.6536
R241C	c.721C>T	7	Missense	0.0300	0.0402	0.0370	0.5453
V399V	c.1197A>T	11	Silence/splice	0.0250	0.0357	0.0321	0.4999
IVS6–1G>A	c.707–1G>A	Intron 6	Splice	0.0207	0.0283	0.0247	0.4349
G257V	c.770G>T	7	Missense	0.0259	0.0189	0.0222	0.4419
Y166X	c.498C>A	5	Nonsense	0.0052	0.0283	0.0173	0.0776
V388M	c.1162G>A	11	Missense	0.0259	0.0094	0.0173	0.1878
S70del	c.208_210delTCT	3	Deletion	0.0104	0.0189	0.0148	0.3878
R241Pfs	c.722delG	7	Deletion	0.0052	0.0236	0.0148	0.1311
W326X	c.977G>A	10	Nonsense	0.0207	0.0094	0.0148	0.2995
IVS7 + 2T>A	c.842+2T>A	Intron 7	Splice	0.0100	0.0893	0.0099	0.6524
R400T	c.1199G>C	11	Missense	0.0052	0.0142	0.0099	0.3476
R400K	c.1199G>A	11	Missense	0.0104	0.0094	0.0099	0.6524
R408Q	c.1223G>A	12	Missense	0.0104	0.0094	0.0099	0.6524
A434D	c.1301C>A	12	Missense	0.0104	0.0094	0.0099	0.6524
R158Q	c.473G>A	5	Missense	0.0104	0.0047	0.0074	0.4648
L242F	c.724C>T	7	Missense	0.0052	0.0094	0.0074	0.5352
L255S	c.764T>C	7	Missense	0.0155	0.0000	0.0074	0.1073
R261Q	c.782G>A	7	Missense	0.0104	0.0047	0.0074	0.4648
A342T	c.1024G>A	10	Missense	0.0155	0.0000	0.0074	0.1073
E56D	c.168G>T	2	Missense	0.0104	0.0000	0.0049	0.2265
IVS4 + 3G>C	c.441+3G>C	Intron 4	Splice	0.0052	0.0047	0.0049	0.7266
Y154H	c.460T>C	5	Missense	0.0000	0.0094	0.0049	0.2856
A156P†	c.466G>C	5	Missense	0.0104	0.0000	0.0049	0.2265
R157K	c.470G>A	5	Missense	0.0052	0.0047	0.0049	0.7266
F161S	c.482T>C	5	Missense	0.0104	0.0000	0.0049	0.2265
R176X	c.526C>T	6	Nonsense	0.0000	0.0094	0.0049	0.2856
Y206C	c.617A>G	6	Missense	0.0052	0.0047	0.0049	0.7266
I224T	c.671T>C	6	Missense	0.0104	0.0000	0.0049	0.2265
G247V	c.740G>T	7	Missense	0.0052	0.0047	0.0049	0.7266
T278I	c.833C>T	7	Missense	0.0000	0.0094	0.0049	0.2856
E280K	c.838G>A	7	Missense	0.0104	0.0000	0.0049	0.2265
S349A	c.1045T>G	10	Missense	0.0000	0.0094	0.0049	0.2856
C357Y†	c.1070G>A	11	Missense	0.0052	0.0047	0.0049	0.7266
F39del	c.115_117delTTC	2	Deletion	0.0000	0.0047	0.0025	0.5349
R53H	c.158G>A	2	Missense	0.0052	0.0000	0.0025	0.4765
IVS2+5G>C	c.168+5G>C	Intron 2	Splice	0.0052	0.0000	0.0025	0.4765
I65T	c.194T>C	3	Missense	0.0000	0.0047	0.0025	0.5349
S70P	c.208T>C	3	Missense	0.0052	0.0000	0.0025	0.4765
F121L†	c.361T>C	4	Missense	0.0000	0.0047	0.0025	0.5349
IVS4+2T>A†	c.441+2T>A	Intron 4	Splice	0.0052	0.0000	0.0025	0.4765
IVS5+1G>A	c.509+1G>A	Intron 5	Splice	0.0052	0.0000	0.0025	0.4765
Y154C†	c.461A>G	5	Missense	0.0000	0.0047	0.0025	0.5349
R158W	c.472C>T	5	Missense	0.0052	0.0000	0.0025	0.4765
Q172X	c.514C>T	6	Nonsense	0.0000	0.0047	0.0025	0.5349
E183G†	c.548A>G	6	Missense	0.0000	0.0047	0.0025	0.5349
L227Q†	c.680T>A	6	Missense	0.0052	0.0000	0.0025	0.4765
E228X†	c.682G>T	6	Nonsense	0.0000	0.0047	0.0025	0.5349
IVS5–2A>G†	c.510–5A>G	Intron 5	Splice	0.0000	0.0047	0.0025	0.5349
G239D	c.716G>A	7	Missense	0.0000	0.0047	0.0025	0.5349
G247R	c.739G>C	7	Missense	0.0052	0.0000	0.0025	0.4765
R252W	c.754C>T	7	Missense	0.0052	0.0000	0.0025	0.4765
R252Q	c.755G>A	7	Missense	0.0000	0.0047	0.0025	0.5349
R261X	c.781C>T	7	Nonsense	0.0000	0.0047	0.0025	0.5349
Q267H	c.801G>C	7	Missense	0.0000	0.0047	0.0025	0.5349
P275A†	c.823C>G	7	Missense	0.0000	0.0047	0.0025	0.5349

(*Continued*)

Table 2. *(Continued)*

Trivial name (protein effect)	Systematic name (DNA level)	Exon	Characters of mutation	Relative frequency(%) Nothern China (n = 100)	Southern China (n = 112)	Chinese (n = 212)	p
M276K	c.827T>A	7	Missense	0.0000	0.0047	0.0025	0.5349
R270G†	c.810A>G	7	Missense	0.0052	0.0000	0.0025	0.4765
I269N	c.806T>A	7	Missense	0.0000	0.0047	0.0025	0.5349
IVS7 + 1G>A	c.842+1G>A	Intron 7	Splice	0.0000	0.0047	0.0025	0.5349
S310F	c.929C>T	9	Missense	0.0000	0.0047	0.0025	0.5349
I324N†	c.971T>A	10	Missense	0.0000	0.0047	0.0025	0.5349
Q336R	c.1007A>G	10	Missense	0.0000	0.0047	0.0025	0.5349
A345T	c.1033G>A	10	Missense	0.0000	0.0047	0.0025	0.5349
G346R	c.1036G>A	10	Missense	0.0000	0.0047	0.0025	0.5349
S350Y	c.1049C>A	10	Missense	0.0000	0.0047	0.0025	0.5349
C357X†	c.1071C>A	11	Nonsense	0.0000	0.0047	0.0025	0.5349
P362L†	c.1085C>T	11	Missense	0.0052	0.0000	0.0025	0.4765
L367R	c.1100T>G	11	Missense	0.0052	0.0000	0.0025	0.4765
S411X	c.1232C>A	12	Nonsense	0.0052	0.0000	0.0025	0.4765
Y414X†	c.1242C>A	12	Nonsense	0.0052	0.0000	0.0025	0.4765
Y414C	c.1241A>G	12	Missense	0.0000	0.0047	0.0025	0.5349
IVS12 + 6T>A	c.1315+6T>A	Intron 12	Splice	0.0000	0.0047	0.0025	0.5349
E286K	c.856G>A	8	Missense	0.0000	0.0047	0.0025	0.5349

* The p value is for the comparison of the relative frequencies between two geographical groups.
† Novel mutation which being not reported in the "PAHdb" at January 8, 2007.

polymorphisms, 15 novel mutations occurred at a low frequency (0.2–0.4%).

Genotype. Complete *PAH* locus genotyping was established for 194 of 212 patients (91.5%). A further 17 had one mutation allele identified; no mutations were detected in one patient. In the group of the fully genotyped patients, 22 were homozygous for one mutation, which is among the most frequent ones in the total pool of Chinese *PAH* mutant alleles while the remaining 172 patients were compound heterozygote for two mutations. The homozygosity value in the Chinese population is only 8.9%, and the genotypic homozygosity frequency is 11.3% (22/194), which are the consequence of the relatively low frequency of prevalent mutations, further pointing to a high heterogeneity of Chinese population. According to our mutation analysis, a total of 127 genotypes were detected, and R243Q × R243Q, R243Q × IVS4 − 1G>A, and R243Q × Ex6–96A>G were dominant with frequencies of 7.3, 4.4, and 2.9%, respectively. We also compared the frequencies of the genotypes between two regions; only two genotypes demonstrated significant differences, they were R111X × R243Q and R243Q × R413P.

DISCUSSION

We have systematically investigated the variety of genetic defects underlying PKU in a sample of 212 patients representing a cross-section of Chinese Han population. A point mutation or a microdeletion was identified in the coding region or immediately adjacent intronic regions of *PAH* gene on 405 of 424 independent chromosomes. Among the 79 different mutations found, 68.4% were missense mutations, 15.2% were splice mutations, 12.7% were nonsense mutations, and 3.8% were frameshift deletions. Eight mutations, with a relative frequency of 3% or more, accounted for two thirds of the identified ones. In this context, a total of 15 previously unknown mutations were identified in the Chinese population, and all of them were found to be rare. The data presented here indicated that the total pool of mutant *PAH* alleles, at least in patients of Chinese descent, consisted of a small number of frequent mutations and a very high number of rare mutations. This distribution of different types of mutation was very similar to that observed in European and other Asian populations (3–5,11).

As revealed by our intriguing findings and useful analyses, we found that:

1. Previous molecular studies during the last 18 y have elucidated the spectrum of mutations in PKU patients of a few Asian populations and indicated that mutations were not randomly distributed and particular ones showed regional associations. Derived from our study, eight mutations, R243Q, Ex6–96A>G, IVS4 − 1G>A, R413P, Y356X, R111X, R241C, and V399V, with a relative frequency of 3% or more, account for 62% of all mutant alleles. Meanwhile, the most prevalent mutations in Japanese PKU alleles are R413P, IVS4 − 1G>A, R241C, R243Q, T278I, Ex6 − 96A>G, Y356X, and R111X, accounting for 74.4% of all (5). In Korean patients (3), the most common mutations are R243Q, IVS4 − 1G>A, and Ex6 − 96A>G, each with a frequency of 10% or more. However, R111X—a frequent mutation in Japanese and Chinese patients—is very rare in Korean, whereas R413P—the most prevalent one in Japanese—forms a very small proportion of patients in Korean. T278I—with a relative high frequency in Japanese and Korean—is very rare (0.5%) in Chinese patients. However, for Y356X, there is a similar relative frequency in Chinese, Japanese, and Korean patients (5.2, 4.9, and 5.7%). In summary, a total of eight different mutations—R243Q, EX6 − 96A>G, R111X, Y356X, R413P, IVS4 − 1G>A, R241C, and T287I—were

identified that might be regarded as prevalent in Asian population, which reached relative frequencies of at least 3% in at least two countries, the same criteria used by Zschocke in his systematic review on *PAH* gene mutation spectrum in Europe (11).

Comparison of the PKU mutational spectrum among Orientals has identified the aggregate of specific prevalent mutation. Of further interesting, it is visible that some rare mutations cluster in particular regions, for instance, 442-706delE5/E6 accounting for 2.4% in Japanese patients (5), E286 K 4% in Taiwanese (4), and A259T 5.7% in Koreans (3). All these characteristics are in agreement with those of mutation profiles and their frequencies varying among populations, with many alleles being specific to regions.

Although the 424 independent chromosomes analyzed here represent only a small fraction of the total pool of mutant PAH alleles in Chinese Han population, the cross-sectional results provide some insight into the extent of mutational heterogeneity. Calculation of the heterogeneity (homozygosity value) at the PAH locus in Chinese as previously described (15) gives a value of 8.9%, which might be in part due to strict discouragement or even prohibition of consanguineous marriages. Moreover, the value is lower than that of other Asian populations or regions, such as Japanese and Taiwanese, indicating the higher heterogeneous the Chinese population with respect to PAH mutation and the gene flow occurring between Asian populations.

2. To map the distribution of PKU mutations in Chinese Han population and unveil the underlying origins and mechanisms, the allele frequencies of our study were stratified by geographic regions, named southern and northern China. We found that except for R413P, which gave a significant p value ($p = 0.0155$) between two regions, with a relative allele frequency of 8% in the northern PKU Chinese but 2.7% in the southern ones, seven other common mutations, including R243Q, EX6 − 96A>G, IVS4 − 1G>A, Y356X, R111X, V399V, and R241C reached at least 3% in both regions without a significant difference ($p > 0.05$) statistically (Table 2).

Furthermore, we compared our data of the northern group with that derived from Song *et al.* (7). A total of 10 mutations were included when they reached a relative allele frequency of at least 3% in either study, R243Q, EX6–96A>G, IVS4 − 1G>A, Y356X, R111X, R241C, R413P, V399V, IVS7 + 2T>A, and R53H. There were no remarkable differences on these common mutations but for R111X ($p = 0.0322$) between two databases, which suggested our sample could be representative of the northern group. Then we combined the alleles of our northern group with that of Song *et al.*, compared the affiliated allele frequencies with that of our southern group, and we observed that only R413P, R111X, and R53H, with a p value of 0.0189, 0.0373, and 0.036 respectively, showed differences between two regions (Table 3). Do these findings contradict the evidence on local mutation clustering determined in previous studies (21–23,28)? Or could we contribute this uniform distribution of common *PAH* deficiency mutations other than R413P and R53H to migration and mixture between regions? We can offer no obvious explanation for these data. The data limitation is a possibility, as existing ones were selective, of small sample size and an ambiguous ethnic background of patients. Migration and admixture between regions or founder effect are not unlikely alternatives. To elucidate the underlying possibility with the data available, we attempted to link R413P and R241C to range expansion and migration of early historic populations before it became mixed in present-day populations. Published manuscripts on R413P in Orientals indicated that it accounted for 30.5% in Japanese PKU alleles (5), 6.5% in northern Chinese (7), 3.2% in Korean (3), and 4% in Taiwanese (4). According to our study, it occurred at a relative allele frequency of 5.4% on average, with 8.3% in north china, and 2.8% in south. These data suggested that this specific allele might spread throughout the Orient by a founder effect, validating the hypothesis that "northern Mongoloids" (35) represented as a founding population in Asia in that PKU mutation might have occurred in northern Mongoloids and subsequently spread to the Chinese and Japanese population. As for R241C, which representing a strong founder effect in Taiwanese (4) who mainly derived from Chinese, has never been described as a common mutation (7,22,26,36). But in our study, it was found that R241C has a relative frequency of 3% or more in both regions. This discrepancy could be contributed to the evidence that there might be multiple founding populations of PKU in East Asia and genetic drift. Furthermore, all the ambiguity presented here led us to doubt on the traditional geographical stratification method (19,20) in the Chinese Han population, because a geographical label that is usually adequate for the overall classification of samples may only represent a certain proportion of the actual underlying population genetic structure, as the real information of human history is hidden in the genome (37,38).

Table 3. *Geographic distribution of the mutant PAH alleles within China. Allele counts for ten common mutations are given together with their relative frequencies expressed as percentage values (in brackets)*

Mutant allele	Northern China Combined*	Southern China Present study	p†
R243Q	133 (23.2)	52 (23.2)	0.9976
EX6 − 96A>G	54 (9.5)	23 (10.3)	0.718
IVS4 − 1G>A	23 (4.0)	14 (6.3)	0.1774
R413P	40 (7.0)	6 (2.7)	0.0189
Y356X	35 (6.1)	10 (4.5)	0.3644
R111X	40 (7.0)	7 (3.1)	0.0373
R241C	10 (1.8)	9 (4.0)	0.0587
V399V	16 (2.8)	8 (3.6)	0.5634
IVS7 + 2T>A	10 (1.8)	2 (0.9)	0.2986
R53H	10 (1.8)	0 (0.0)	0.036
Other	199	93	
Total	570	224	

* Allele frequencies summed for the northern Chinese of the data from our study and that from Song *et al.*7

† The p value is for the comparison of the relative frequencies between the combined northern Chinese and southern ones from our study.

3. Derived from our study, the geographical and ethnic source of the common mutations in Chinese could not be estab-

lished, and it would be interesting to expand this study to an investigation of haplotypes to clarify their origins and also assess genetic admixture, the effect of migration, expansion, and founder effect. We hope to find these answers in research yet to set up.

REFERENCES

1. Guttler F 1980 Hyperphenylalaninemia: diagnosis and classification of the various types of phenylalanine hydroxylase deficiency in childhood. Acta Paediatr Scand Suppl 280:1–80
2. Scriver CR, Eisensmith RC, Woo SL 1995 The hyperphenylalaninemias. In: Scriver CR, Beaudet AL, Sly WS, Valle D (eds) The Metabolic and Molecular Bases of Inherited Disease, 7th ed. McGraw-Hill, New York, pp 60
3. Lee DH, Koo SK, Lee KS, Yeon YJ, Oh HJ, Kim SW, Lee SJ, Kim SS, Lee JE, Jo I, Jung SC 2004 The molecular basis of phenylketonuria in Koreans. J Hum Genet 49:617–621
4. Chien YH, Chiang SC, Huang A, Chou SP, Tseng SS, Huang YT, Hwu WL 2004 Mutation spectrum in Taiwanese patients with phenylalanine hydroxylase deficiency and a founder effect for the R241C mutation. Hum Mutat 23:206–212
5. Okano Y, Asada M, Kang Y, Nishi Y, Hase Y, Oura T, Isshiki G 1998 Molecular characterization of phenylketonuria in Japanese patients. Hum Genet 103:613–618
6. Okano Y, Hase Y, Lee DH, Furuyama J, Shintaku H, Oura T, Isshiki G 1992 Frequency and distribution of phenylketonuric mutations in Orientals. Hum Mutat 1:216–220
7. Song F, Qu YJ, Zhang T, Jin YW, Wang H, Zheng XY 2005 Phenylketonuria mutations in Northern China. Mol Genet Metab 86:S107–S118
8. Daniele A, Cardillo G, Pennino C, Carbone MT, Scognamiglio D, Correra A, Pignero A, Castaldo G, Salvatore F 2007 Molecular epidemiology of phenylalanine hydroxylase deficiency in Southern Italy: a 96% detection rate with ten novel mutations. Ann Hum Genet 71:185–193
9. Stojiljkovic M, Jovanovic J, Djordjevic M, Grkovic S, Cvorkov Drazic M, Petrucev B, Tosic N, Karan Djurasevic T, Stojanov L, Pavlovic S 2006 Molecular and phenotypic characteristics of patients with phenylketonuria in Serbia and Montenegro. Clin Genet 70:151–155
10. Zschocke J, Preusse A, Sarnavka V, Fumic K, Mardesic D, Hoffmann GF, Baric I 2003 The molecular basis of phenylalanine hydroxylase deficiency in Croatia. Hum Mutat 21:399–405
11. Zschocke J 2003 Phenylketonuria mutations in Europe. Hum Mutat 21:345–356
12. Santana da Silva LC, Carvalho TS, da Silva FB, Morari L, Fachel AA, Pires R, Refosco LF, Desnick RJ, Giugliani R, Saraiva Pereira ML 2003 Molecular characterization of phenylketonuria in South Brazil. Mol Genet Metab 79:17–24
13. Kasnauskiene J, Giannattasio S, Lattanzio P, Cimbalistiene L, Kucinskas V 2003 The molecular basis of phenylketonuria in Lithuania. Hum Mutat 21:398–403
14. Zschocke J, Hoffmann GF 1999 Phenylketonuria mutations in Germany. Hum Genet 104:390–398
15. Guldberg P, Levy HL, Hanley WB, Koch R, Matalon R, Rouse BM, Trefz F, de la Cruz F, Henriksen KF, Guttler F 1996 Phenylalanine hydroxylase gene mutations in the United States: report from the Maternal PKU Collaborative Study. Am J Hum Genet 59:84–94
16. Eiken HG, Knappskog PM, Boman H, Thune KS, Kaada G, Motzfeldt K, Apold J 1996 Relative frequency, heterogeneity and geographic clustering of PKU mutations in Norway. Eur J Hum Genet 4:205–213
17. Guldberg P, Romano V, Ceratto N, Bosco P, Ciuna M, Indelicato A, Mollica F, Meli C, Giovannini M, Riva E, Biasuccl G, Henriksen KF, Guttler F 1993 Mutational spectrum of phenylalanine hydroxylase deficiency in Sicily: implications for diagnosis of hyperphenylalaninemia in southern Europe. Hum Mol Genet 2:1703–1707
18. Avigad S, Cohen BE, Bauer S, Schwartz G, Frydman M, Woo SL, Niny Y, Shiloh Y 1990 A single origin of phenylketonuria in Yemenite Jews. Nature 344:168–170
19. Jin L, Su B 2000 Natives or immigrants: modern human origin in east Asia. Nat Rev Genet 1:126–133
20. Chu JY, Huang W, Kuang SQ, Wang JM, Xu JJ, Chu ZT, Yang ZQ, Lin KQ, Li P, Wu M, Geng ZC, Tan CC, Du RF, Jin L 1998 Genetic relationship of populations in China. Proc Natl Acad Sci USA 95:11763–11768
21. He WP, Lai M, Zhang QZ, Zhang XS, Li XL, Ying SG 1997 Study on mutation diversity of phenylalanine hydroxylase gene in two groups of Chinese and one group of Japanese. Chin J Med Genet 14:344–347
22. Zhang M, Gu XF, Zhang MH, Zhang YF, Pan XS, Huang XD, Shen YN, Ye J, Chen RG 1995 Mutations and their frequencies in exon 7 of phenylalanine hydroxylase gene of Phenylketonuria in southern Chinese. Chin J Med Genet 12:324–327
23. Lo WH, Wang T, Eisensmith R, Woo SL 1993 Molecular basis of PKU in China. Chin Med Sci J 8:180–185
24. Song F, Wu GY, Xu GZ, Cai WN, Ding XY 1995 Frequency of five point mutations of phenylalanine hydroxylase and prenatal gene diagnosis of phenylketonuria. Chin J Med Genet 12:321–324
25. Song F, Jin YW, Wang H, Yang YL, Zhang YM, Zhang T 2003 Ten novel mutations in the phenylalanine hydroxylase gene identified in Chinese patients with phenylketonuria. Zhongguo Yi Xue Ke Xue Yuan Xue Bao 25:142–144
26. Song F, Jin YW, Wang H, Zhang YM, Yang YL, Zhang T 2005 Mutations in exon 7 of the phenylalanine hydroxylase (PAH) gene in chinese patients with phenylketonuria. Yi Chuan 27:53–56
27. Qiu WJ, Zhang YF, Ye J, Han LS, Gu XF 2004 [Study on mutations of exon 12 of the PAH gene in 127 phenylketonuria patients]. Zhonghua Yi Xue Yi Chuan Xue Za Zhi 21:261–263
28. Yu WM, Xu L, Li XW, He C, Shen M, Zhang ZX, Jin YY, Zhou ZS, Qiao F 2003 An eighteen-year study on phenylketonuria. Zhongguo Yi Xue Ke Xue Yuan Xue Bao 25:218–222
29. Yu WZ, Qiu DH, Song F, Liu L, Jin MW, He J, Gui JH, Wang R, Zhou HY, Wang Z, Zhou Y, Lei Q, Zhang ZP, Liu XW 2007 Studies on mutations of exon 11 and 12 in phenylalaninase gene of Phenylketonuria patients in Xinjiang. Med J Chin PLA 32:326–328
30. Wang N, Zhu Y, Xu K, Qiu Z, Song W, Huang S 1999 [Identification of mutations in the phenylalanine hydroxylase gene and exon 5 novel mutation Y166X(C–>G) in Yunnan]. Zhonghua Yi Xue Yi Chuan Xue Za Zhi 16:9–11
31. John SW, Weitzner G, Rozen R, Scriver CR 1991 A rapid procedure for extracting genomic DNA from leukocytes. Nucleic Acids Res 19:408
32. Nowacki P, Byck S, Prevost L, Scriver CR 1996 The PAH mutation analysis consortium database: update. Nucleic Acids Res 25:139–142, 1997
33. Stenson PD, Ball EV, Mort M, Phillips AD, Shiel JA, Thomas NS, Abeysinghe S, Krawczak M, Cooper DN 2003 Human gene mutation database (HGMD):update. Hum Mutat 21:577–581, 2003
34. Cotton RG, Scriver CR 1998 Proof of "disease causing" mutation. Hum Mutat 12:1–3
35. Matsumoto H 1988 Characteristics of Mongoloid and neighboring populations based on the genetic markers of human immunoglobulins. Hum Genet 80:207–218
36. Gu XF, Zhang M, Chen RG 1995 Phenylketonuria mutations in southern Chinese detected by denaturing gradient gel electrophoresis in exon 7 of PAH gene. J Inherit Metab Dis 18:753–754
37. Shi Y, Zhao X, Yu L, Tao R, Tang J, La Y, Duan Y, Gao B, Gu N, Xu Y, Feng G, Zhu S, Liu H, Salter H, He L 2004 Genetic structure adds power to detect schizophrenia susceptibility at SLIT3 in the Chinese Han population. Genome Res 14:1345–1349
38. Pritchard JK, Stephens M, Rosenberg NA, Donnelly P 2000 Association mapping in structured populations. Am J Hum Genet 67:170–181