

Phenylketonuria mutations in Northern China

Fang Song^{a,*}, Yu-jin Qu^a, Ting Zhang^a, Yu-wei Jin^a, Hong Wang^a, Xiao-ying Zheng^{b,*}

^a Department of Medical Genetics, Capital Institute of Pediatrics, Beijing, PR China

^b Institute of Population Research, Peking University, Beijing, PR China

Received 17 June 2005; received in revised form 26 August 2005; accepted 1 September 2005

Available online 26 October 2005

Abstract

Mutation spectrum of phenylalanine hydroxylase (*PAH*) gene in patients with phenylketonuria (PKU) in Northern China is described with a discussion on genotype–phenotype correlation. By using PCR/SSCP and DNA sequencing, all exons of *PAH* gene in the 185 unrelated patients with PKU from Northern China were studied. A total of 70 different mutations, including 42 missense, 12 splice, 7 nonsense, 5 deletion, 3 insertion, and 1 silence/splice mutations, were detected in 349/370 mutant alleles (94.3%). Deletion, insertion, and frameshift mutations were found for the first time in China PKU patients. The mutations R243Q, EX6-96A>G, R111X, Y356X, and R413P were the prevalent mutations with relative frequencies of 22.2, 11.1, 8.7, 6.5, and 6.5%, respectively. Fifteen novel mutations were identified in this study: I38fsX19, IVS4+3G>C, Y154H, R157K, R157I, T200fsX6, Q267H, Q267E, F302fsX39, G346R, S349A, L367L, R400K, IVS12+4A>G, and IVS12+6T>A. Each of them occurs at very low frequency (0.3–1.1%). The mutation spectrum of PKU in Chinese is similar to other Asian populations but significantly different from European populations. Altogether, 70 different mutations are found in 109 genotypes distributed among 185 PKU patients. As shown by the analysis, the predicted residual activity found in the majority of PKU individuals match their *in vivo* phenotypes, though evidence is also found for both phenotypic inconsistencies among subjects with similar genotypes and discordance between the *in vitro* and *in vivo* effects of some mutant alleles. The study enables us to construct a national database in China serving as a valuable tool for genetic counseling and prognostic evaluation of future cases of PKU.

© 2005 Elsevier Inc. All rights reserved.

Keywords: Phenylketonuria; Phenylalanine hydroxylase; Mutation detection; Genotype–phenotype correlation; China

Introduction

Phenylketonuria (PKU; MIM 261600) is the most common autosomal recessive disorder in China. The incidence of PKU as determined by newborn mass screening in Chinese is 1/11,000 [1], which is similar to that in European (1/10,000) [2], but higher than that in Japanese (1/120,000) [3], and Korean (1/41,000) [4]. PKU is caused by a deficiency of hepatic phenylalanine hydroxylase (PAH). The disease causes hyperphenylalaninaemia (HPA) and severe mental retardation unless the affected child is early put on a strict low-phenylalanine diet [5]. The associated phenotypes due

to PAH deficiency range in severity from classical PKU through mild PKU to MHP, as are defined by pretreatment blood phenylalanine levels.

Since the identification of the *PAH* gene in 1986 [6], almost 500 different mutations have been identified and listed in the *PAH* mutation database (PAHdb; <http://www.mcgill.ca/pahdb>). Over the last years, comprehensive mutation data have become available for most European and a few Asian countries. There are marked differences in the spectrum of mutations and in the degree of heterogeneity between different populations, but in many countries, there are a few common mutations that account for the majority of alleles in PKU patients. Although, there have been some reports describing *PAH* gene mutations in Chinese populations [7–11], the mutation analysis was limited to a few common mutations or exons; the information on mutation spectrum of Chinese PKU is incomplete.

* Corresponding authors. Fax: +86 10 85628367 (F. Song); +86 10 62751976 (X. Zheng).

E-mail addresses: songf_558@263.net (F. Song), xzheng@pku.edu.cn (X. Zheng).

Recent studies of *PAH* 3D structures [12] reveal that the *PAH* enzyme is comprised of four monomeric proteins. Each monomer has three structural domains; a N-terminal regulatory domain (residues 1–142), a catalytic domain (residues 143–410), and a C-terminal tetramerization domain (residues 411–452). The studies have provided information on active site and the binding sites for its substrate and cofactor.

Mutation analysis of a given population is useful for further understanding of the structural and functional aspects of the mutant protein and the correlation between genotype and phenotype. It is helpful in facilitating genetic consultation of patients' families. In this paper, we present a comprehensive analysis of mutations in 370 independent PKU alleles from 185 patients in Northern China, collate the mutation relative frequencies in some Asian [13,4] and European populations [14–18], and analyze the genotype–phenotype correlation.

Subjects and methods

Subjects

Patients

A total of 185 patients with *PAH* deficiency in 185 independent families from Northern China are included in this study. Most of the patients (135) were identified when they showed mental retardation between 6 months and 3 years old. Only 50 patients were identified in neonatal screening. *PAH* deficiency was diagnosed by conventional biochemical methods. Patients are assigned to three phenotype categories according to the pretreatment plasma phenylalanine (Phe) levels: classic PKU (Phe \geq 1200 $\mu\text{mol/L}$), mild PKU (Phe 600–1200 $\mu\text{mol/L}$), and MHP (Phe $<$ 600 $\mu\text{mol/L}$). Parental samples were available from a part of the families. There is no consanguinity among the families studied. Informed consent for genetic analysis was obtained from the parents of the patients.

Control

Fifty apparently healthy children (100 *PAH* alleles) with a blood concentration Phe $<$ 120 $\mu\text{mol/L}$ were sampled as the control for mutation analysis.

Mutation analysis

Genomic DNA was extracted from leucocytes of patients and both parents (when available) according to standard protocols and screened for DNA alterations. Eight exons (2, 3, 5, 6, 7, 10, 11, and 12) of *PAH* gene were first screened by SSCP before the samples showing abnormal SSCP migration patterns were sequenced to confirm the mutations. Those alleles where no mutation was found were sequenced for all the eight exons and their flanking intron regions so that no mutation therein might be missed. The exons 1, 4, 8, 9, and 13 were sequenced directly. Maternal or paternal inheritance of mutations was determined

particularly for patients with novel mutation when parental DNA was available. Mutations detected in the *PAH* gene were identified by comparison with the *PAH* cDNA sequence which was based on the GenBank reference sequence (NM_000277.1; GI: 4557818). Mutation nomenclature was referred to the websites at <http://www.hgvs.org/mutnomen> and <http://www.pahdb.mcgill.ca>. Novel mutations were defined by exclusion from the PAHdb (PAHdb; <http://www.mcgill.ca/pahdb>) and previously reported mutations on PubMed (<http://www.ncbi.nlm.nih.gov/PubMed>).

PCR amplification

The 13 exons and their exon-flanking intronic sequences of *PAH* gene were amplified by PCR. The sequences of primer for PCR were designed in accordance with the literature, without the CG clamp at the 5' end [19]. Reaction mixtures were made in a volume of 25 μl containing 0.5 μg of genomic DNA, 2–4 pmol of each primer, 0.25 mM of each dNTP, and the reaction buffer supplied with the *Taq* DNA polymerase (Hua Mei Biotechnology). The PCR cycles were carried out at 97 $^{\circ}\text{C}$ for 5 min for initial denaturing, then 30 cycles at 94 $^{\circ}\text{C}$ for 45 s, at 55 $^{\circ}\text{C}$ for 45 s and at 72 $^{\circ}\text{C}$ for 45 s, followed by a final extension of 8 min at 72 $^{\circ}\text{C}$ in a Biometra Thermocycler.

Single-strand conformation polymorphism analysis

Five microliters of each PCR product was mixed with 5 μl of denaturation solution (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, and 0.05% xylene cyanol) and heated for 5 min at 97 $^{\circ}\text{C}$, then cooled on ice immediately. Samples were separated on 6–10% non-denaturing PAGE gel (49:1). For electrophoresis of the different fragments, the conditions were varied within the following ranges: 30–60 W/600 V, 3.5–5.5 h, 10–28 $^{\circ}\text{C}$, for 20 cm gel. The gel was stained with silver nitrate for visualizing the fragment migration.

Sequence analysis

Sequence analysis was carried out directly with PCR products. Domestic commercial companies, AUGCT and BioAsia, which use ABI 377 automatic sequencer, performed sequencing.

Genotype–phenotype analysis

In the analysis, nonsense mutation, frameshift mutation, and splice-site mutation are counted as null mutations and three phenotypes are distinguished: classical PKU, mild PKU, and MHP.

For our analysis, the genotypes are first divided into two categories: homoallelic mutant *PAH* genotype and heteroallelic mutant *PAH* genotype. Next, the heteroallelic *PAH* genotypes are further listed roughly in the order of increasing predicted residual activity (PRA), showing a transition from Null + Null through Null + Missense (functionally heterozygous) and finally to Missense + Missense.

Results

Identification of the *PAH* mutations

Mutations were identified on 349 of 370 PKU alleles in the present study, representing a mutation detection rate of 94.3%. The relative frequency of all mutations identified in the *PAH* gene in Northern China patients with HPA is shown in Table 1.

Seventy distinct mutations were identified, including missense (42/70), splice (12/70), nonsense (7/70), deletion (5/70), insertion (3/70), and silence/splice (1/70) mutations. Three insertion (I38fsX19, T200fsX6, and L367fsX27) and two deletion (R241fsX5 and F302fsX39) mutations caused frameshift, leading to premature termination of transcription. The other three deletion mutations are in-frame, the result of loss of 3 bp for a code of *PAH* gene.

The most prevalent mutations in Chinese PKU are R243Q, EX6-96A>G, R111X, Y356X, and R413P, representing 22.2, 11.1, 8.7, 6.5, and 6.5% of the 370 mutant alleles, respectively. The next common mutations are IVS4-1G>A, V399V, R53H, and IVS7+2T>A, each of them showing a relative frequency over 2.0%.

In this series of cases, exon 7 and its flanking introns contain the greatest number of mutant alleles (123/370, 33.2%). The other exons and their flanking intronic sequences with mutant alleles, listed from high to low mutation frequencies, are: exon 6 (49/370), 11 (49/370), 3 (40/370), 12 (31/370), 5 (29/370), 2 (12/370), 4 (6/370), 10 (5/370), 8 (3/370), and 9 (2/370). No mutation is detected in exons 1 and 13. Thirteen mutations occur at CpG dinucleotide motifs.

Among the 70 mutations, two are found near common polymorphism sites, leading to the appearance of two cDNA changes in each case. At codon 232 of *PAH*, a mutation results in a transition C>T at c.694 while a neighboring polymorphism gives rise to another transition A>G at c.696. These two changes turned the codon CAA into a stop one TAG, resulting in premature termination of transcription. Similarly, in addition to the mutational deletion of G at c. 722, a neighboring polymorphism results in a transversion G>A at c.735, again turning the codon for the 5th amino acid after the deletion into a stop codon.

Novel mutations

In total, 28 mutations have not been fully described previously in the literature. Among them, four mutations were briefed in a preliminary report in 2003 [20]: G239D, R241fsX5, R400T, and IVS11+2T>C; nine were submitted to the PAHdb on May 17, 2004: Y206C, Q235P, G247S, G247R, E280G, D282G, Q336R, A395D, and A434D.

Fifteen novel mutations are reported in this paper: I38fsX19, IVS4+3G>C, Y154H, R157K, R157L, T200fsX6, Q267H, Q267E, F302fsX39, G346R, S349A, L367L, R400K, IVS12+4A>G, and IVS12+6T>A.

Each of the above mutations occurs at low frequency (0.3–1.1%). And they are not found in screening of 100 controls *PAH* alleles for each.

Genotype and phenotype correlation

Among the 185 patients, 152 are classical PKU; 22, mild PKU; 1, MHP while, for the remaining 10 cases, information about their clinical phenotype is not available (Table 2). As a result, analysis into genotype–phenotype correlation is performed mainly on the 175 patients with available information.

Among the 185 patients, the 70 different mutations are variously combined into 109 genotypes. In our study, nonsense mutation, frameshift mutation, and splice-site mutation are counted as null mutations. Since the 23 kinds of null mutation involve more than 40% PKU alleles (158/370) (Table 1), the genotypes are divided in the way as described in Mutation analysis: Homoallelic mutant *PAH* genotype (7) and Heteroallelic mutant *PAH* genotype (102), the latter of which is further divided into Null + Null (19), Null + Missense (52), and Missense + Missense (31). The majority of patients are compound homozygotes (165/185, 89%); among them, 87 are functional hemizygotes (Null + Missense). Through analysis of the homozygous, the Null + Null, and the functional heterozygous genotypes, we have confirmed the effects that 63 of the 70 mutations exert on in vivo phenotypes. The majority (48) of mutations produce definite and consistent effects: 43 cause classical PKU; and 5, mild PKU. Inconsistencies, however, are found in the following two aspects: (1) 11 mutations appear in two phenotypes: 9 mutations (S70del, IVS4-1G>A, EX6-96A>G, R243Q, Y356X, P362T, R400T, R413P, A434D) appear in both classical and mild PKU, R241C in both mild PKU and MHP, and V399V in both classical PKU and MHP (i.e. wide apart in classification). (2) 4 mutations show discordance between in vitro (unit protein) and in vivo phenotypes. Although the enzyme activities expressed in vitro of F39del, R261Q, I65T, and I95del are 20, 30, 21, and 27%, respectively, all of them behave like null mutations, giving rise, when combined with other null mutations, to classical PKU.

R111X, IVS4-1G>A, EX6-96A>G, R243Q, Y356X, V399V, and R413P are common mutations in PKU patients in China. R111X, IVS4-1G>A, EX6-96A>G, Y356X, V399V, and R413P have been proven as being null mutations [21–23] and the PRA value for R243Q is <10%. Yet, except for R111X which shows concordance between in vitro and in vivo phenotypes, discordance is noted in all the remaining mutations in either homoallelic mutant *PAH* genotype or heteroallelic *PAH* genotype. Nevertheless, excluding from account the 10 patients with no information available, all these mutations show fairly high correlation with the classical PKU: IVS4-1G>A, 10/11; EX6-96A>G, 35/41; R243Q, 68/80; Y356X, 20/23; V399V, 9/10; and R413P, 18/21.

Table 1
Spectrum of PAH mutations identified in northern Chinese

Trivial name (protein effect)	Systematic name (DNA level)	Location	Characters of mutation	PKU1		PKU2		Total of PKU		Novel mutation	CpG
				No. of alleles	RF(%)	No. of alleles	RF(%)	No. of alleles	RF(%)		
I38fsX19	c.111_112insG	Exon 2	Insertion	1	0.4			1	0.3	*	
F39del	c.115_117delTTC	Exon 2	Deletion			2	2.0	2	0.5		
R53H	c.158G>A	Exon 2	Missense	8	3.0	1	1.0	9	2.4		
I65T	c.194T>C	Exon 3	Missense	1	0.4	1	1.0	2	0.5		
S70del	c.208_210delTCT	Exon 3	Deletion	3	1.1	1	1.0	4	1.1		
I95del	c.284_286delTCA	Exon 3	Deletion	2	0.7			2	0.5		
R111X	c.331C>T	Exon 3	Nonsense	27	10.0	5	5.0	32	8.7		#
IVS4+1G>A	c.441+1G>A	Intron 4	Splicing	2	0.7			2	0.5		
IVS4+3G>C	c.441+3G>C	Intron 4	Splicing	4	1.5			4	1.1	*	
IVS4-1G>A	c.442-1G>A	Intron 4	Splicing	10	3.7	1	1.0	11	3.0		
Y154H	c.460T>C	Exon 5	Missense			2	2.0	2	0.5	*	
R157K	c.470G>A	Exon 5	Missense	1	0.4			1	0.3	*	
R157I	c.470G>T	Exon 5	Missense	1	0.4			1	0.3	*	
R158Q	c.473G>A	Exon 5	Missense	3	1.1			3	0.8		#
F161S	c.482T>C	Exon 5	Missense	3	1.1	1	1.0	4	1.1		
Y166X	c.498C>G	Exon 5	Nonsense	6	2.2	1	1.0	7	1.9		
R176X	c.526C>T	Exon 6	Nonsense	1	0.4			1	0.3		#
Y179H	c.535T>C	Exon 6	Missense	1	0.4			1	0.3		
T200fsX6	c.598_599insA	Exon 6	Insertion			1	1.0	1	0.3	*	
EX6-96A>G	c.611A>G	Exon 6	Splicing	30	11.0	11	11.0	41	11.1		
Y206C	c.617A>G	Exon 6	Missense	1	0.4			1	0.3		
Y206X	c.618C>A	Exon 6	Nonsense	1	0.4			1	0.3		
I224T	c.671T>C	Exon 6	Missense	1	0.4			1	0.3		
Q232X	c.694C>T;c.696A>G	Exon 6	Nonsense	1	0.4			1	0.3		
Q235P	c.704A>C	Exon 6	Missense	1	0.4			1	0.3		
G239D	c.716G>A	Exon 7	Missense	1	0.4			1	0.3		
R241C	c.721C>T	Exon 7	Missense	2	0.7	2	2.0	4	1.1		#
R241fsX5	c.722delG; c.735G>A	Exon 7	Deletion	1	0.4	2	2.0	3	0.8		#
R243Q	c.728G>A	Exon 7	Missense	62	23.0	20	20.0	82	22.2		#
G247S	c.739G>A	Exon 7	Missense			1	1.0	1	0.3		
G247R	c.739G>C	Exon 7	Missense			1	1.0	1	0.3		
G247V	c.740G>T	Exon 7	Missense	5	1.9			5	1.4		
L249H	c.746T>A	Exon 7	Missense			1	1.0	1	0.3		
R252Q	c.755G>A	Exon 7	Missense	3	1.1	3	3.0	6	1.6		#
L255S	c.764T>C	Exon 7	Missense	1	0.4			1	0.3		
R261Q	c.782G>A	Exon 7	Missense	1	0.4	2	2.0	3	0.8		#
Q267H	c.801G>C	Exon 7	Missense			1	1.0	1	0.3	*	
Q267E	c.799C>G	Exon 7	Missense	1	0.4			1	0.3	*	
M276K	c.827T>A	Exon 7	Missense			1	1.0	1	0.3		
E280G	c.839A>G	Exon 7	Missense	1	0.4			1	0.3		
E280K	c.838G>A	Exon 7	Missense	1	0.4			1	0.3		
P281L	c.842C>T	Exon 7	Missense	1	0.4			1	0.3		#
IVS7+1G>A	c.842+1G>A	Intron 7	Splicing	1	0.4			1	0.3		#
IVS7+2T>A	c.842+2T>A	Intron 7	Splicing	8	3.0			8	2.3		
D282G	c.845A>G	Exon 8	Missense	1	0.4			1	0.3		
F302fsX39	c.904delT	Exon 8	Deletion	1	0.4			1	0.3	*	
S303P	c.907T>C	Exon 8	Missense	1	0.4			1	0.3		
IVS8-7A>G	c.913-7A>G	Intron 8	Splicing	1	0.4			1	0.3		
A309D	c.926G>A	Exon 9	Missense			1	1.0	1	0.3		
W326X	c.977G>A	Exon 10	Nonsense	1	0.4			1	0.3		
Q336R	C.1007A>G	Exon 10	Missense			1	1.0	1	0.3		
G346R	c.1036G>A	Exon 10	Missense			1	1.0	1	0.3	*	
G346R	c.1036G>C	Exon 10	Missense	1	0.4			1	0.3		
S349A	c.1045T>G	Exon 10	Missense			1	1.0	1	0.3	*	
IVS10-11G>A	c.1066-11G>A	Intron 10	Splicing	1	0.4	1	1.0	2	0.5		
Y356X	c.1068C>A	Exon 11	Nonsense	15	5.6	9	9.0	24	6.5		
P362T	c.1084C>A	Exon 11	Missense			2	2.0	2	0.5		
L367fsX27	c.1099_1100insC	Exon 11	Insertion	1	0.4			1	0.3		
L367L	c.1101G>A	Exon 11	Silence/Splicing			1	1.0	1	0.3	*	
A395D	c.1184C>A	Exon 11	Missense	1	0.4			1	0.3		
V399V	c.1197A>T	Exon 11	Splicing	8	3.0	3	3.0	11	3.0		
R400K	c.1199G>A	Exon 11	Missense	2	0.7			2	0.5	*	

Table 1 (continued)

Trivial name (protein effect)	Systematic name (DNA level)	Location	Characters of mutation	PKU1		PKU2		Total of PKU		Novel mutation	CpG
				No. of alleles	RF(%)	No. of alleles	RF(%)	No. of alleles	RF(%)		
R400T	c.1199G>C	Exon 11	Missense	4	1.5			4	1.1		
IVS11+2T>C	c.1199+2T>C	Intron 11	Splicing	1	0.4			1	0.3		
R408W	c.1222C>T	Exon 12	Missense	1	0.4			1	0.3		#
R413P	c.1238G>C	Exon 12	Missense	19	7.0	5	5.0	24	6.5		#
T418P	c.1252A>C	Exon 12	Missense	1	0.4			1	0.3		
A434D	c.1301C>A	Exon 12	Missense			3	3.0	3	0.8		
IVS12+4A>G	c.1315+4A>G	Intron 12	Splicing			1	1.0	1	0.3	*	
IVS12+6T>A	c.1315+6T>A	Intron 12	Splicing			1	1.0	1	0.3	*	
Unknown				258	95.6	91	91.0	349	94.3		
Total				12	4.4	9	9.0	21	5.7		
				270	100.0	100	100.0	370	100.0		

PKU1, symptomatic patients; PKU2, patients recognized by neonatal screening; RF, relative frequency; *, novel mutation; #, CpG sites.

Discussion

Need for gene analysis in China

PKU is a common metabolic disorder in China, which can be detected by neonatal screening and treated with low-phenylalanine diet. Since we have achieved coverage of only 10% in neonatal screening [1], however, most PKU patients so far detected were diagnosed only after clinical symptoms were noted. It is a good news for us that the nationwide coverage for neonatal screening has shown an annual increase of 45.5% for the recent 6 years, and a coverage about 98% has been achieved in a few large cities [1]. That is to say, more and more patients are being identified at an early time after birth. As a result, more information on the genetics of PKU is needed so that the PKU families might be better able to early detect and early initiate effective intervention on the patients, as well as, to take effective preventive measures. Hopefully, analysis of *PAH* mutations and correlation study of the genotype and phenotype will prove helpful to genetic counseling for them.

Profile of the *PAH* mutations

A mutation detection rate of 94.3% of PKU alleles is achieved in 185 Chinese PKU patients with 70 different mutations. Missense mutations account for the majority of mutations (42/70) with a relative frequency of 49.2% (182/370). The small deletion and insertion mutations are identified for the first time in Chinese PKU population. Mutations are distributed over most coding sequence except exons 1 and 13, but occur mainly in the highly conserved regions spanning exons 3, 5–7, and 11–12. The high degree of heterogeneity in *PAH* gene mutation results in most of patients being compound heterozygous (165/185). Some reports on PKU in China [7–11] were confined to screening relatively small patient cohorts for a small number of selected mutations. The present study extends these previous results to a more comprehensive understanding of *PAH* allele distribution and frequency in Chinese.

Most samples reported in this paper are patients of classical PKU that were detected after presentation of symptoms. This is certainly a result of the present limited coverage in neonatal screening.

The frequencies of mutations detected after presentation of symptoms and that of mutations detected in screening are listed separately in Table 1. As shown by comparison of five common mutations, the two frequencies are near each other for EX6-96A>G, R243Q, and R413P, but the frequency of mutations found in symptomatic patients is greater for R111X while the frequency of mutations detected in screening is greater for Y356X. As for the frequencies found with other mutations, the values obtained are too low for statistical analysis. All the patients detected in screening were provided by the Beijing Newborn Disease Screening Center. Although Beijing is one of the cities with the greatest population mobility, most of the mobile populations come from the northern provinces. In consequence, the frequency distribution of common *PAH* mutations as found by screening in Beijing regions comes near that of northern provinces.

No mutation could be detected in 21 alleles despite scanning of the entire *PAH* coding region by sequencing. The most likely explanation is that the mutations are located in regions undetectable in this study, i.e., in the 5' and 3' untranslated regions, in the intronic sequences far away from exons, or the mutations may be large deletions. However, the technique used in this study could not detect large deletion mutations. So far these mutant types have not been reported in Chinese PKU, but they are reported in other countries though at a rather low relative frequency [13].

Comparison of relative frequencies of common *PAH* gene mutations

It is well known that different ethnic groups have their own distinctive and diverse *PAH* mutant allele series including one or a few prevalent founder alleles. For collating the relative frequencies of *PAH* mutant among Asian and European populations, we have consulted Zschocke's

Table 2
Genotype–phenotype association in 185 Chinese PKU patients

Classification	Genotypes	PRA (%)	Number of patients			
			Classical	Mild	MHP	NA
Homozygous	R111X + R111X	Null	4			
	IVS4-1G>A + IVS4-1G>A	Null	1			
	EX6-96A>G + EX6-96A>G	Null	2		1	
	R243Q + R243Q	<10	9			
	F39del + F39del	20	1			
	R261Q + R261Q	30	1			
	Y154H + Y154H	?	1			
Heterozygous	R111X + IVS4-1G>A	Null + Null	1			
	R111X + Y166X	Null + Null	1			
	R111X + EX6-96A>G	Null + Null	3			
	R111X + IVS7+2T>A	Null + Null	2			
	R111X + VSI10-11G>A	Null + Null	2			
	R111X + Y356X	Null + Null	1			
	R111X + V399V	Null + Null	1			
	Y166X + IVS4+1G>A	Null + Null	1			
	Y166X + EX6-96A>G	Null + Null	1			
	Y166X + Y356X	Null + Null	1			
	Y166X + V399V	Null + Null	1			
	EX6-96A>G + IVS4+3G>A	Null + Null	1			
	EX6-96A>G + Y206X	Null + Null	1			
	EX6-96A>G + R241fsX5	Null + Null	1			
	EX6-96A>G + W326X	Null + Null	1			
	EX6-96A>G + R356X	Null + Null	2			
	EX6-96A>G + V399V	Null + Null	2			
	R241fsX5 + IVS12+4A>G	Null + Null	1			
	Y356X + IVS4-1G>A	Null + Null	3			
	P281L + IVS7+2T>A	<1 + Null	1			
	R408W + Y356X	<1 + Null	1			
	E280K + IVS4+3 G>C	1-3 + Null				1
	R413P + R111X	<3 + Null	3			
	L255S + IVS4+1G>A	<3 + Null	1			
	R413P + IVS4+3 G>C	<3 + Null				1
	R413P + EX6-96A>G	<3 + Null	2		1	
	R413P + IVS7+1G>A	<3 + Null	1			
	R413P + IVS7+2T>A	<3 + Null	1			
	R413P + Y356X	<3 + Null	1			
	R413P + V399V	<3 + Null	1			1
	R252Q + V399V	3 + Null	1			
	R252Q + EX6-96A>G	3 + Null	1			
	G247V + Y204C	4 + Null	1			
	G247V + I8-7 a>g	4 + Null	1			
	G247V + IVS4+3G>C	4 + Null				1
	F161S + EX6-96A>G	7 + Null	2			
	R243Q + R111X	<10 + Null	5			
	R243Q + IVS4-1G>A	<10 + Null	2		1	
	R243Q + Y166X	<10 + Null	2			
	R243Q + EX6-96A>G	<10 + Null	7		1	
	R243Q + R241fsX5	<10 + Null	1			
	R243Q + Q232X	<10 + Null				1
	R243Q + IVS7+2T>A	<10 + Null	2			
R243Q + Y356X	<10 + Null	7		1		
R243Q + L367fsX27	<10 + Null	1				
R243Q + V399V	<10 + Null	1				
I65T + R111X	21 + Null	1				
R241C + V399V	25 + Null				1	
I95del + Y356X	27 + Null	2				
R261Q + IVS7+2T>A	30 + Null	1				
S70del + F302fsX39	? + Null	1				
S70del + R111X	? + Null	1				
Y179H + R176X	? + Null				1	
Y206C + EX6-96A>G	? + Null			1		
Q235P + IVS4-1G>A	? + Null	1				

Table 2 (continued)

Classification	Genotypes	PRA (%)	Number of patients			
			Classical	Mild	MHP	NA
	G239D + V399V	? + Null	1			
	G247S + T 200fsX6	? + Null	1			
	Q267E + R111X	? + Null	1			
	A309D + Y356X	? + Null	1			
	G346R + EX6-96A>G	? + Null	1			
	S349A + EX6-96A>G	? + Null	1			
	P362T + V399V	? + Null	1			
	R400T + IVS7+2T>A	? + Null	1			
	T418P + EX6-96A>G	? + Null	1			
	A434D + Y356X	? + Null		1		
	? + I38fsX19	? + Null	1			
	? + R111X	? + Null	1			1
	? + IVS4-1G>A	? + Null	1			
	? + EX6-96A>G	? + Null	3	1		
	? + Y356X	? + Null	1	1		1
	? + IVS11+2T>C	? + Null	1			
	R413P + R243Q	<3 + <10	4	1		1
	R413P + G247R	<3 + ?	1			
	R413P + IVS12+6 T>A	<3 + ?		1		
	R413P + ?	<3 + ?	4			
	R252Q + F161S	3 + 7	1			
	R252Q + R243Q	3 + <10	1			
	R252Q + Q267H	3 + ?	1			
	R252Q + A434D	3 + ?	1			
	G247V + F161S	4 + 7	1			
	G247V + ?	4 + ?	1			
	R158Q + E280G	10 + ?	1			
	R158Q + R400K	10 + ?				1
	R243Q + R158Q	<10 + 10	1			
	R243Q + I65T	<10 + 21	1			
	R243Q + R241C	<10 + 25		3		
	R243Q + S70del	<10 + ?		1		
	R243Q + R157K	<10 + ?	1			
	R243Q + I224T	<10 + ?	1			
	R243Q + L249H	<10 + ?		1		
	R243Q + Q336R	<10 + ?	1			
	R243Q + L367L	<10 + ?	1			
	R243Q + R400K	<10 + ?	1			
	R243Q + R400T	<10 + ?	1	1		
	R243Q + A434D	<10 + ?	1			
	R243Q + ?	<10 + ?	8	2		
	S70del + P362T	? + ?		1		
	D282G + S303P	? + ?	1			
	G346R + R400T	? + ?	1			
	R157I + ?	? + ?		1		
	M276K + ?	? + ?		1		
	A395D + ?	? + ?	1			
Total	109		152	22	1	10

Null: frameshift mutation, Splice-site mutations and Nonsense mutations; **PRA:** predicted residual activity in vitro, which comes from the web: <http://pahdb.mcgill.ca>; **Phenotypes:** Classic = classical PKU; Mild = mild PKU; MHP = mild hyperphenylalaninaemia; NA = Not available.

systematic review on *PAH* gene mutations in Europe [24], where mutations with a relative allele frequency of at least 3% (totaling 29) were included. Similarly, 11 mutations with relative frequency >3% are chosen from our series for comparison. As there are differences among the European populations themselves, geological and ethnic factors are taken into account as well and populations representing different common European *PAH* mutants are selected for comparison with our data (Table 3).

Comparison of the mutations of PKU in Asia reveals several significant differences, as well as some overlaps of mutant allele distribution among Chinese, Japanese, and Korean populations. R243Q and EX6-96A>G, two of the most frequent mutations in Chinese, are also frequently detected in Korean and Japanese. However, R111X, a frequent mutation in Chinese and Japanese patients, is very rare in Korean patients. The relative frequency of Y356X is very similar in Chinese, Japanese, and Korean (6.5,

Table 3
Relative frequencies (%) of common *PAH* gene mutation detected in Asian and European populations

Geographic distribution	Mutations	CpG mutation	Northern Chinese 1:11,000	Japanese [13] 1:120,000	Korean [4] 1:41,000	Dane [14] 1:12,000	German [15] 1:8,000	Czech [16] 1:9,000	Spain [17] (Catalonia) 1:6,600	Irishman [18] 1:4,500
Asia	R111X	#	9	4	<1	<1	0	0	1	0
	IVS4-1G>A		3	7	10	0	0	0	0	0
	EX6-96A>G		11	6	10	0	0	0	0	0
	R241C (MHP)	#	1	7	6	0	<1	0	0	0
	R243Q	#	22	7	12	0	1	1	3	0
	A259T		0	0	6	0	0	0	0	0
	T278I		0	7	3	0	0	0	0	0
	Y325X		0	0	3	0	0	0	0	0
	Y356X		7	5	6	<1	0	0	0	0
	V399V		3	0	0	0	0	0	0	0
	R413P	#	7	31	3	0	0	0	0	0
	Europe	F39L		0	0	0	1	<1	2	0
G46S			0	0	0	1	<1	0	0	<1
L48S			0	0	0	0	3	2	<1	0
I65T			<1	0	0	<1	<1	<1	6	10
A104D			0	0	0	3	<1	<1	0	<1
R158Q		#	<1	0	<1	3	4	5	2	<1
W187X			0	0	0	0	<1	0	0	0
R243X		#	0	0	0	0	1	3	2	3
R252W		#	0	0	0	0	<1	2	<1	<1
R261Q		#	<1	0	<1	2	6	2	6	<1
R261X		#	0	0	<1	0	<1	0	3	0
G272X			0	0	0	2	<1	3	0	0
E280K		#	<1	0	0	3	<1	0	<1	1
P281L		#	<1	0	2	1	4	<1	0	<1
IVS7+1G>A		#	<1	0	0	<1	<1	<1	<1	<1
F299C			0	0	0	0	<1	<1	0	4
A300S(MHP)			0	0	0	0	<1	0	3	0
L348V			0	0	0	2	<1	0	0	5
S349P			0	0	0	<1	<1	0	5	1
IVS10-11G>A			<1	0	0	5	10	3	13	1
c.1089DelG			0	0	0	0	0	0	0	0
V388M		#	0	1	3	<1	0	<1	5	0
E390G			0	0	0	0	2	0	3	<1
A403V(MHP)			0	0	0	0	2	0	5	0
R408W -H2		#	<1*	0	0	17	22	55	3	0
R408W-H1		#	0	0	0	1	0	0	0	41
R408Q		#	0	1	<1	<1	<1	0	0	<1
Y414C			0	0	0	10	6	2	1	1
IVS12+1G>A			0	0	0	37	10	4	<1	2
Total				94	93	86	99	98	95	90
Number of alleles			370	82	136	308	546	266	198	558

* The R408W mutation in our results was not subjected to haplotype analysis and was placed randomly at position R408W-H2.

4.9, and 5.7%). Yet, R413P, the most prevalent allele in Japanese, forms a very small proportion of patients in Korean. In contrast, IVS4-1G>A occupies a relatively larger proportion in Korean mutant allele profiles than in Chinese or Japanese. Besides, the mutations T278I and R241C have relative high frequencies in Japanese and Korean, but for Chinese, it is either very low (R241C) or not detected (T278I).

The frequencies of common mutations of Asian and European are listed in Table 3. R243Q, EX6-96A>G, R111X, Y356X, R413P, and IVS4-1G>A appear to be the most prevalent mutations in Asian populations. On the other hand, there are several prevalent founder alleles in the European populations: R408W is the most

common *PAH* mutation, followed by IVS10-11G>A, IVS12+1G>A, Y414C, and I65T. Indeed, remarkable differences are noted between the mutation spectrum of Asian and that of European. In addition, differences are also noted in mild mutations. While R241C is the predominant type in the Asian spectrum, A300S and A403V make the greater part in European. Between the Asian and European populations, great differences are frequent, both with regard to kind and relative frequencies, suggesting entirely different origins of the *PKU* gene mutations between them. Although the same mutations may be found between them in some CpG sites, the hotspots for *PKU* mutation, yet significant differences are noted between their relative frequencies (see Table 3). It is thus suggested that the

mutations occurring at these CpG sites are recurrences. However, it requires further haplotype study for confirmation.

Novel mutations

Fifteen novel mutations plus four that were briefed in a preliminary report in 2003 [20] and 9 that were submitted to PAHdb in May 17, 2004, 28 mutations in total were studied together in this paper. Their absence in 100 normal *PAH* alleles for each has ruled out the possibility of their being polymorphisms. Family gene analysis accomplished for each of them has also confirmed that the allele was inherited from the parent who did not carry other PKU mutations.

Novel mutations are defined by reference to the data published at PAHdb or on PubMed on June 3, 2005. But these databases are mainly based on materials derived from studies among European. It is thus no wonder that so many novel mutations are detected in this comprehensive analysis of PKU mutations in Northern China.

Frameshift

Among the 28 mutations, 4 mutations (I38fsX19, T200fsX6, R241fsX5, and F302fsX39) cause frameshift that results in the premature termination of translation. The *PAH* proteins translated from such mRNAs lost the C-terminal 396, 247, 207, and 142 amino acids, respectively, which contain the entire catalytic domain and the tetramerization domain (I38fsX19) or a part of the catalytic domain and the entire tetramerization domain (T200fsX6, R241fsX5, and F302fsX39) [12]. Thus, both the function and stability of the truncated proteins are severely affected. It is highly likely that these mutations may be the disease-causing mutations.

Mutations affecting RNA processing

Six mutations (IVS4+3G>C, IVS11+2T>C, IVS12+4A>G, IVS12+6T>A, R400K, and R400T) in the 5' splice-site consensus sequence are identified. The sequence between -2 and +6 of the splice-donor site has been regarded as highly conserved [25], and any mutation involving these sites is considered likely to affect the correct splicing of the intron. Ketterling et al. have postulated that splice-donor mutations not involving the invariant GT doublet at the splice-donor junction would result in abolishment of more than 99% of normal splicing if the splice-donor sequence matched the consensus in five or six bases of the eight-base consensus sequence [26]. In our series, except for IVS11+2T>C, which directly involved the invariant GT doublet at the splice-donor junction of intron 11, the mutations IVS4+3G>C, IVS12+4A>G, IVS12+6T>A, R400K (c.1199G>A), and R400T (c.1199G>C) affect the splice-donor sites of *PAH* introns 4, 12, and 11 by matching the eight-base consensus sequence in five or six bases (Fig. 1).

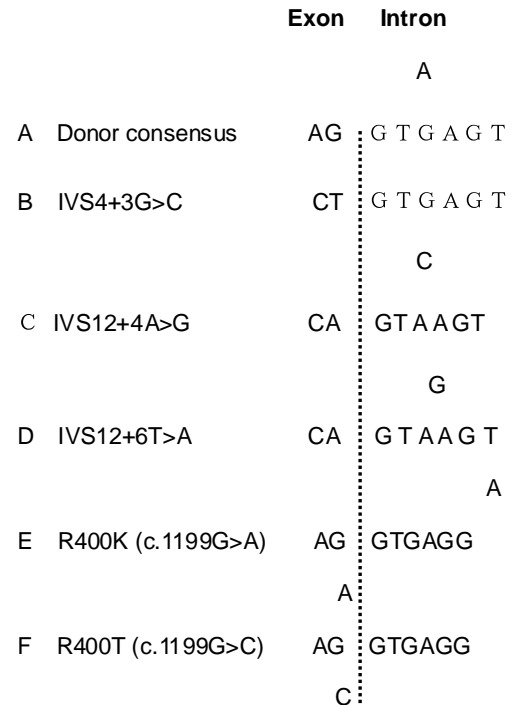


Fig. 1. The mutations affecting the consensus splice-donor sequence. (A) The consensus splice-donor sequence. (B) The c.441+3 G>C substitution at nucleotide position +3 of the *PAH* intron 4. (C) The c.1315+4 A>G substitution at nucleotide position +4 of the *PAH* intron 12. (D) The c.1315+6 T>A substitution at nucleotide position +6 of the *PAH* intron 12. (E) The c.1199 G>A substitution at nucleotide position -1 of the *PAH* intron 11. (F) The c.1199 G>C substitution at nucleotide position -1 of the *PAH* intron 11.

To determine whether IVS12+4A>G and IVS12+6T>A are disease-causing mutations or mere polymorphisms, we performed splicing score calculation based on the method of Shapiro and Senapathy [27]. The splicing score for the original 5' donor sequence of the 12th exon of *PAH* gene is 79.4. But after the two mutations, IVS12+4A>G and IVS12+6T>A, the splicing scores are reduced to 68.6 and 73.7, respectively. Among them, the mutation IVS12+4A>G causes a reduction of 11.3. It may thus be speculated that these two mutations might have lowered the splicing efficiency or led to frank failure of correct splicing of intron 12. However, we have no direct evidence to corroborate our speculation and further mRNA study is needed for confirmation.

L367L is associated with classical PKU, though it is a silent mutation in the 11th exon of *PAH* gene. Silent mutations have been reported to cause diseases mainly by affecting RNA processing. For example, several alleles, including Q304Q (c.912G>A) [28] and V399V (c.1197A>T) [23], initially called silent, were later found to promote exon skipping. There is also evidence indicating that silent mutation in exons might cause disease by disrupting exonic splicing enhancer (ESE). ESEs are present in constitutive or alternative exons of certain genes, and are required for efficient splicing on those exons. The ESEs in pre-mRNAs are recognized by serine/arginine-rich (SR) proteins, a

family of essential splicing factors that regulate alternative splicing [29,30]. Using the motif-scoring matrices in ESEfinder release 2.0 (<http://exon.cshl.edu>), we tried to find high-score motifs in exon 11 of wild-type and mutant PAH gene for four SR proteins (SF2/ASF, SC35, SRp40, and SRp35). Multiple high-score motifs of each type are found to be distributed throughout this exon. L367L mutation specifically disrupts the third (CCCCTGG) of seven high-score SRp40 motifs. The score for the third high-score SRp40 motif is 3.3264 in wild-type PAH and is reduced to 2.7586 in L367L mutation PAH. Surprisingly, at the same time, L367L mutation produced a new high-score SRp40 motif (CTAGAGC, with a score of 3.0887), which is not found in the normal PAH sequence.

In this connection, Liu's study on BRCA1, a breast cancer susceptibility gene, might be cited [30], in which a nonsense mutation in codon 18 reduced the score of a high-score SF2/ASF in wild-type gene from 2.134 to 0.079 in the mutant gene and experiment revealed exon skipping caused by the mutation. Liu et al. also analyzed a database of 50 single-base substitutions in human genes that cause exon skipping *in vivo*, and found that more than one-half of the single-base substitutions reduced or eliminated at least one high-score motif for one or more of these SR proteins. In our study, while reducing the score of a SRp40 motif, however, L367L also produced a new high-score SRp40 motif. Using the SR protein scoring matrices, it is still difficult to determine whether this situation affects correct splicing by disrupting exonic splicing enhancer. It appears that RNA-level study is required to explore the effect of L367L on PAH protein function.

Missense mutations

The remaining 17 mutations are missense mutations. Sixteen of them are located in the catalytic domain (residues 143–410) and the remaining one, in the tetramerization domain (residues 411–452) of PAH protein. As judged from structural information on PAH protein [12], the amino acid Y154, R157, Y206, S349, and A395 are relevant to the normal structure of PAH monomer; E280 and D282 have a bearing on the shape of the active center of the enzyme; A434 directly affects the stability of the PAH protein tetramer. As a result, mutations at these sites may affect the activity of the protein by altering the structure or stability of the protein. Based on our genotype–phenotype correlation study, Y154H, R157K, S349A, A395D, E280G, and D282G lead to classical PKU while Y206C and A434D give rise to mild PKU. There are however reports that missense mutation in exons may also affect the enzyme functions by affecting RNA splicing [21]. For example, although no marked reduction in enzyme activity has been demonstrated in *in vitro* expression study of EX6-96A>G, it has been shown that a new splicing site is produced that causes exon skipping. It is therefore important that, for newly discovered missense mutations, in-depth functional study is needed to determine whether they are disease causing and to learn their pathogenic mechanisms.

Genotype–phenotype correlation

Even though the PKU mutations found in China are highly heterogeneous and the genotypes of most patients are compound heterozygotes, yet classical PKU makes up the predominant type in our sample. The presence of null mutation in 40% of alleles and the availability of PRA data for many mutations have proved helpful to our genotype–phenotype correlation study.

Studying genotype–phenotype correlation in homoallelic mutant PAH genotype and in Null + Null and Null + Missense (functionally heterozygous) genotypes have enabled us to discover the effect that single mutation exerts on phenotype. Equipped with the knowledge thus gained, we tried next to study compound heterozygotes. As shown by the results, there does exist wide-ranging matching between PAH mutant genotype and PKU phenotype. In most of the cases, the effects are severe, including most of the mutations discussed under the section Novel Mutations. This may be explained by the high proportion presence of null mutations in our sample.

In the cases of I157K, Y206C, R241C, L249H, M276K, and IVS12+6, most of them have no records on their enzyme activity expressed *in vitro* except for R241C (PRA = 25%). Yet the fact that their combination with severe mutations produced mild clinical phenotypes has led us to regard them as mild mutations. Just as Guldberg et al. [31] put it, “disease severity in most cases is determined by the least severe of two PAH mutations.” As for R241C, although one of the four patients containing R241C is of the genotype R241C+V399V and manifests MHP phenotype, yet all the rest are of the genotype R241C+R243Q and manifests mild PKU. Again this led us to regard it as a mild mutation, a result in keeping with that reported in Taiwan [32].

Discordance between *in vitro* and *in vivo* phenotypes

While the observed metabolic phenotypes of F161S, G247V, R252Q, L255S, E280K, P281L, and R408W match their predicted gene effects, inconsistency is found with F39del, I95del, I65T, and R261Q. The inconsistency with I65T and R261Q has long been reported. In contrast to *in vitro* expression tests the result of which is a unit protein, the consequence of a single mutation, most PKU patients are heterozygotes comprising different PAH mutations. So far, it is still a knotty problem to sort out the complicated interactions among the different subunits of PAH protein produced by different mutations. This is really one of the reasons for this discordance.

Interindividual inconsistency in *in vivo* PAH mutation effect

Interindividual inconsistency is manifested in the common PAH mutations: IVS4-1G>A, EX6-96A>G, R243Q, Y356X, V399V, and R413P. As most of them are closely associated with classical phenotype (85–91%), however, they are assigned to the severe type. Inconsistency is also observed in S70del, P362T, R400T, and A434D. Among

them, R400T and A434D are mutations discussed under the section Novel Mutations. Since, information about their enzyme activity expressed in vitro is not available, however, it is difficult to predict their effect on phenotype. Interindividual inconsistency has frequently been reported and its mechanisms intensively discussed in literature [31,33]. Guldberg et al. have postulated that a potential cause of these inconsistencies may be related to the biological properties and functions of the mutant protein.

The result of our study supports the view that the genotype of the PAH gene mutations is the major determinant of the metabolic phenotype of PAH protein deficiency. Yet the actual in vivo expression of the mutation combination has proved difficult to predict. Further in-depth study is needed for elucidating the inconsistency between genotype and phenotype.

Summary

This study is undertaken to produce a detailed picture of the mutation spectrum underlying HPA and PKU in northern Chinese population. The Chinese mutation profile of *PAH* is similar to those of the neighboring Asian populations but significantly different from European population. The definition of the mutational profile of PKU in China enables us to construct a national database covering detailed information on genotype–phenotype correlations. This database would serve as a valuable tool in genetic counseling and in prognostic evaluation of future cases of PKU.

Acknowledgments

This work is supported by research grants from the National Program for Key Basic Research Projects of China (2001CB510306), the Beijing Natural Sciences Foundation (5052008), and the key special foundation of Ministry of Science and Technology of China (2004BA720A04). We are grateful to Dr. Y.L. Yang (Department of Pediatrics, Peking University First Hospital, Beijing, China) and Dr. Y.M. Zhang (Beijing Newborn Disease Screening Center, Beijing, China) for their kind help, and to the families of the PKU patient for their cooperation in providing blood samples and information. We thank Professor Guanyun Wu (Institute of Basic Medical Sciences, CAMS and PUMC, Beijing, China) and Dr. Ju-hsian Chüan (Encyclopedia of China Publishing House, Beijing, China) for helpful discussion and for critical reading of the manuscript.

References

- [1] (in Chinese) X.F. Gu, Zh.G. Wang, Screening for phenylketonuria and congenital hypothyroidism in 5.8 million neonates in China, *China J. Prev. Med.* 2 (2004) 99–102.
- [2] H. Bickel, C. Bachmann, R. Beckers, N.J. Brandt, B.E. Clayton, G. Corrado, H.J. Feingold, O. Giardini, G. Hammersen, D. Schönberg, Neonatal mass screening for metabolic disorders, *Eur. J. Pediatr.* 137 (1981) 133–139.
- [3] K. Aoki, Y. Wada, Outcome of the patients detected by newborn screening in Japan, *Acta Paediatr. Jpn.* 30 (1988) 429–434.
- [4] D.H. Lee, S.K. Koo, K.S. Lee, Y.J. Yeon, H.J. Oh, S.W. Kim, S.J. Lee, S.S. Kim, J.E. Lee, I. Jo, S.C. Jung, The molecular basis of phenylketonuria in Koreans, *J. Hum. Genet.* 49 (2004) 617–621.
- [5] C.R. Scriver, S. Kaufmen, R.C. Eisensmith, S.L.C. Woo, The hyperphenylalaninemias, Seventh ed., in: C.R. Scriver, A.L. Beaudet, W.S. Sly, D. Valle (Eds.), *The Metabolic and Molecular Basis of Inherited Disease*, vol 1, McGraw-Hill, New York, 1995, pp. 1015–1077.
- [6] A.G. Dilella, S.C.M. Kwok, F.D. Ledley, J. Marvit, S.L.C. Woo, Molecular structure and polymorphic map of the human phenylalanine hydroxylase gene, *Biochemistry* 25 (1986) 743–749.
- [7] (in Chinese) F. Song, G.Y. Wu, G.Zh. Xu, W.N. Cai, X.Y. Ding, Frequency of five point mutations of phenylalanine hydroxylase and prenatal gene diagnosis of phenylketuria, *China J. Med. Genet.* 12 (1995) 321–323.
- [8] (in Chinese) M. Zhang, X.F. Gu, M.H. Zhang, Y.F. Zhang, X.Sh. Pan, X.D. Huang, Y.N. Shen, J. Ye, R.G. Chen, Mutations and their frequencies in exon 7 of phenylalanine hydroxylase gene of phenylketonuria in southern Chinese, *China J. Med. Genet.* 12 (1995) 324–326.
- [9] (in Chinese) M. Wang, B.L. Fang, L.F. Yuan, Detection of the site-mutation frequency of phenylalanine hydroxylase and prenatal gene diagnosis of phenylketonuria, *Chin. J. Med.* 72 (1992) 670–673.
- [10] (in Chinese) G.F. Sun, L. Jiang, X. Zhang, B.Zh. Tong, G.Zh. Dong, K.L. Sun, Novel mutations identified in exon 7 of phenylalanine hydroxylase gene in northern Chinese, *Acta Genet. Sinica* 24 (1997) 492–495.
- [11] (in Chinese) L. Song, Y.T. Meng, W.Y. Gao, F.D. Xu, Zh.M. San, Studies on mutation in exon 7 of phenylalanine hydroxylase gene in TianJing Area, *China J. Med. Genet.* 18 (2001) 233–234.
- [12] H. Erlandsen, R.C. Stevens, The structural basis of phenylketonuria, *Mol. Genet. Metab.* 68 (1999) 103–125.
- [13] Y. Okano, M. Asada, Y. Kang, Y. Nishi, Y. Hase, T. Oura, G. Isshiki, Molecular characterization of phenylketonuria in Japanese patients, *Hum. Genet.* 103 (1998) 613–618.
- [14] P. Guldberg, K.F. Henriksen, F. Güttler, Molecular analysis of phenylketonuria in Denmark: 99% of the mutations detected by denaturing gradient gel electrophoresis, *Genomics* 17 (1993) 141–146.
- [15] J. Zschocke, G.F. Hoffmann, Phenylketonuria mutations in Germany, *Hum. Genet.* 104 (1999) 390–398.
- [16] L. Kozák, M. Blažková, V. Kuhrová, A. Pijáčková, Š. Růžicková, S. Štátná, Mutation and haplotype analysis of phenylalanine hydroxylase alleles in classical PKU patients from the Czech Republic: identification of four novel mutations, *J. Med. Genet.* 34 (1997) 893–898.
- [17] J. Mallolas, M.A. Vilaseca, J. Campistol, N. Lambruschini, F.J. Cambra, X. Estivill, M. Milá, Mutational spectrum of phenylalanine hydroxylase deficiency in the population resident in Catalonia: genotype–phenotype correlation, *Hum. Genet.* 105 (1999) 468–473.
- [18] K.A. O'Donnell, C. O'Neill, O. Tighe, G. Bertorelle, E. Naughten, P.D. Mayne, D.T. Croke, The mutation spectrum of hyperphenylalaninaemia in the Republic of Ireland: the population history of the Irish revisited, *Eur. J. Hum. Genet.* 10 (2002) 530–538.
- [19] P. Guldberg, V. Romano, N. Ceratto, P. Ciuna, A. Indelicato, F. Mollica, C. Meili, M. Giovannini, E. Riva, G. Biasucci, K.F. Henriksen, F. Güttler, Mutational spectrum of phenylalanine hydroxylase deficiency in Sicily: implications for diagnosis of hyperphenylalaninemia in Southern Europe, *Hum. Mol. Genet.* 2 (1993) 1703–1707.
- [20] F. Song, Y.W. Jin, H. Wang, Y.L. Yang, Y.M. Zhang, T. Zhang, Ten novel mutations in phenylalanine hydroxylase gene identified in Chinese patients with phenylketonuria, *Acta Acad. Med. Sinica* 25 (2003) 142–144.
- [21] S. Ellingsen, P.M. Knappskog, H.G. Eiken, Phenylketonuria splice mutation (EXON6nt-96A>G) masquerading as missense mutation(Y204C), *Hum. Mutat.* 9 (1997) 88–90.

- [22] S. Ellingsen, P.M. Knappskog, J. Apold, H.G. Eiken, Divers PAH transcripts in lymphocytes of PKU patients with putative nonsense (G272X, Y356X) and missense (P281L, R208Q) mutations, *FEBS Lett.* 457 (1999) 505–508.
- [23] H.K. Chao, K.J. Hsiao, T.S. Su, A silent mutation induces exon skipping in the phenylalanine hydroxylase gene in phenylketonuria, *Hum. Genet.* 108 (2001) 14–19.
- [24] J. Zschocke, Phenylketonuria Mutations in Europe, *Hum. Mutat.* 21 (2003) 345–356.
- [25] S.M. Mount, A catalogue of splice junction sequences, *Nucleic Acids Res.* 10 (1982) 459–472.
- [26] R.P. Ketterling, J.B. Drost, W.A. Scaringe, D.Z. Liao, J.Z. Liu, C.K. Kasper, S.S. Sommer, Reported in vivo splice-site mutations in the factor IX gene: severity of splicing defects and a hypothesis for predicting deleterious splice donor mutations, *Hum. Mutat.* 13 (1999) 221–231.
- [27] M.B. Shapiro, P. Senapathy, RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression, *Nucleic Acids Res.* 17 (1987) 7155–7174.
- [28] P. Guldberg, H.L. Levy, W.B. Hanley, R. Koch, R. Matalon, B.M. Rouse, F. Trefz, F. De la Cruze, K.F. Henriksen, F. Guttler, Phenylalanine hydroxylase gene mutations in the United States: report from the maternal PKU collaborative study, *Am. J. Hum. Genet.* 59 (1996) 84–94.
- [29] B.J. Blencowe, Exonic splicing enhancers: mechanism of action, diversity and role on human genetic diseases, *TIBS* 25 (2000) 106–110.
- [30] H.X. Liu, L. Cartegni, M.Q. Zhang, A.R. Krainer, A mechanism for exon skipping caused by nonsense or missense mutations in BRCA1 and other genes, *Nat. Genet.* 27 (2001) 55–58.
- [31] P. Guldberg, F. Rey, J. Zschocke, V. Romano, B. François, L. Michiels, K. Ullrich, G.F. Hoffmann, P. Burgard, H. Schmidt, C. Meli, E. Riva, I. Dianzani, A. Ponzzone, J. Rey, F. Güttler, A European multicenter study of phenylalanine hydroxylase deficiency: classification of 105 mutations and a general system for genotype-based prediction of metabolic phenotype, *Am. J. Genet.* 63 (1998) 71–79.
- [32] Y.H. Chien, S.C. Chiang, A. Huang, S.P. Chou, S.S. Tseng, Y.T. Huang, W.L. Hwu, Mutation spectrum in Taiwanese patients with phenylalanine hydroxylase deficiency and a founder effect for the R241C mutation, *Hum. Mutat.* 23 (2004) 206.
- [33] E. Kayaalp, E. Treacy, P.J. Waters, S. Byck, P. Nowacki, C.R. Scriver, Human phenylalanine hydroxylase mutations and hyperphenylalaninemia phenotypes: A metaanalysis of genotype–phenotype correlation, *Am. J. Hum. Genet.* 61 (1997) 1309–1317.