Prenatal diagnosis of Chinese families with phenylketonuria

N. Liu[1], X.D. Kong[1], D.H. Zhao[2], Q.H. Wu[1], X.L. Li[2], H.F. Guo[4], L.X. Cui[4], M. Jiang[1] and H.R. Shi[3]

[1]Prenatal Diagnosis Center, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, China
[2]Neonatal Screening Center, The Third Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, China
[3]Department of Obstetrics and Gynecology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, China
[4]Neonatal Screening Center, Zhengzhou Maternity and Child Care Hospital, Zhengzhou, Henan, China

Corresponding author: H.R. Shi
E-mail: shihuirong_l@163.com

Genet. Mol. Res. 14 (4): 14615-14628 (2015)
Received June 23, 2015
Accepted October 7, 2015
Published November 18, 2015
DOI http://dx.doi.org/10.4238/2015.November.18.25

ABSTRACT. The aim of this study is to investigate the ability to prenatally diagnose phenylketonuria (PKU) by using phenylalanine hydroxylase (PAH) gene mutation analysis combined with short tandem repeat (STR) linkage analysis in 118 fetuses from 112 Chinese families. Genomic DNA was extracted from the peripheral blood from members of 112 families and the exons and exon-intron boundaries of the *PAH* gene were amplified by PCR. PCR products were analyzed by bi-directional Sanger sequencing and multiplex ligation-dependent probe amplification (MLPA). The three variable number of tandem repeat (VNTR) markers PAH-1, PAH-26, PAH-32 were used in the prenatal diagnosis for the PKU families. We identified a spectrum of 63 different mutations, including 61 point mutations and indels, two large exon deletion mutations, and five novel mutations. A substantial proportion of mutant alleles were accounted for by p.R243Q (15.62%), EX6-96AG (9.82%), p.V399V (7.59%), p.Y356X (6.70%), and p.R413P

(5.36%). The same mutations were identified in 31 prenatally genotyped fetuses. We identified 58 fetuses that carried only one mutant allele and 29 fetuses that carried no mutations of PAH and were presumed normal. *PAH* gene mutation analysis combined with STR linkage analysis can provide rapid and accurate prenatal diagnosis for PKU families.

Key words: Phenylketonuria; Phenylalanine hydroxylase gene; Mutation; Prenatal diagnosis

INTRODUCTION

Phenylketonuria (PKU; MIM 261600), the most common inherited disorder of amino acid metabolism, is autosomal recessive and caused by mutations of the phenylalanine hydroxylase (PAH) gene. The overall incidence of PKU in China is 1/11,000 (Gu and Wang, 2004). Loss or absence of PAH activity results in increased concentrations of phenylalanine in the blood and toxic concentrations in the brain. Untreated phenylketonuria is associated with progressive intellectual impairment and can be associated with eczematous rash, autism, seizures, and motor deficits (Blau et al., 2010). Developmental problems, aberrant behavior, and psychiatric symptoms often become apparent as the child grows. Although PKU is a curable inherited disease and an early phenylalanine restricted diet results in 90% of patients having normal mental development, the expensive cost, unstable curative effect, and poor treatment adherence cause some children to suffer from the disease. Prenatal diagnosis is the most effective method to avoid PKU-caused complications. However, a prenatal diagnosis cannot be made from sampling the PAH enzyme in the amniotic fluid as the enzyme is only expressed by the liver (Woo et al., 1983). Therefore, prenatal diagnosis of PKU is feasible only by molecular studies. Cloned human *PAH* gene allows prenatal diagnosis by DNA analysis to deduce the phenotype of the fetus. Currently there are two genotype assays. The first, an alternative approach to mutation analysis, is by linkage analysis of short tandem repeat (STR) polymorphic markers associated with the *PAH* gene locus in intron 3 (Goltsov et al., 1993). This method is mainly used in early prenatal diagnosis but requires that the affected child is available and that a firm diagnosis of PKU has been established. Secondly, Sanger sequencing and multiplex ligation-dependent probe amplification (MLPA) based on PCR can be used. For our study, we used Sanger sequencing and MLPA for gene analysis combined with STR linkage analysis, thereby providing the prenatal diagnosis of 118 fetuses from 112 families.

MATERIAL AND METHODS

Subjects and methods

This study was performed at the First Affiliated Hospital of Zhengzhou University (Zhengzhou, China), was approved by the hospital's Ethics Committee, and was performed according to the principles of the Declaration of Helsinki. All subjects gave their written informed consent.

Study subjects

The 112 families analyzed in this study were recruited from the Genetic Counseling clinic of the Prenatal Diagnosis Center of the First Affiliated Hospital and the Third Affiliated Hospital

of Zhengzhou University from January 2009 to March 2014. One couple was a consanguineous marriage (cousins). All the probands in the research had abnormal blood phenylalanine (Phe) concentrations (>120 μM) and had differential diagnosis detected by BH4-loading test, excluding BH4 deficiency. Five probands were deceased, and thus, their parents were used in place. Five families had twins; thus, we tested 118 fetuses from 112 PKU families.

Collection of samples and genomic DNA extraction

Peripheral blood samples were collected from the patients and parents in 107 families and from the parents in the five families in which the proband was deceased. Ethylenediaminetetraacetic acid (EDTA) was used as an anti-coagulant. Transabdominal chorionic villus sampling, amniocentesis, and umbilical cord blood puncture were performed at different times during gestation to evaluate the at-risk fetus. Genomic DNA was isolated from each sample using a commercial kit (TIANamp DNA Kit, Tiangen Biotech, Beijing, China).

PCR and DNA sequencing

PCR primers were designed by Primer 5.0. The primers covered the sequences of all 13 coding exons of the *PAH* gene (Table 1). PCR was performed using the GeneAmp PCR System 9700 (Applied Biosystems, Foster City, CA). The reaction included 20 to 50 ng genomic DNA, 1 μL each primer, 13 μL 2X Taq PCR MasterMix (contains dNTPs, Tris-HCl, Taq polymerase, KCl, and $MgCl_2$), and ddH_2O to a final volume of 25 μL (Taq PCR Mastermix (KT201), Tiangen Biotech, Beijing, China). The PCR amplification was as follows: 5 min initial denaturation at 96°C, followed by 35 cycles of denaturation for 30 sec at 96°C, annealing for 40 s at 58 to 64°C (depending on the primer), and extension for 1 min at 72°C for 1 min, followed by a final extension for 7 min at 72°C. PCR products were evaluated by 1% agarose gel electrophoresis.

Table 1. Primer sequences of *PAH* gene.

Exon	Primer sequence (5′→3′)	Length of PCR products(bp)
E1	F: ATTTGAAGACCACTGGCCAAA R: GAGTCCCGGAAGTGCCTAAAC	656
E2	F: TTGCTTTGTCCATGGAGGTTT R: ACAGGATCTGGAACAGGCAGA	251
E3	F: TGTGAACTAACTGCCCCACCT R: TTGCTGTTATTTTGCCCAAGC	495
E4	F: GGGATCCCCACTTCTGATCTC R: AACAACTCTGCCAACTGCAGG	505
E5	F: CCCCCATTCAAAGCATTCATA R: CATTCAAGATTTCAGGCCAGC	543
E6	F: CCCTTTCATGTGGGAAATCAA R: GTGCTTGTAGGAATGCATGCA	481
E7	F: ATGTCCCTGGGCAGTTATGTG R: TGAGAACAGGAACAAGTGGCA	512
E8	F: GGGAGCATGTCCACAGGAATA R: TATGATCCCACCTGAAATGGG	470
E9	F: GGCCACCCATCACCTTTTTAT R: GTAGCCCTTGAAAACCCTTGG	387
E10	F: TCCCTTCATCCAGTCAAGGTG R: ATTCCAAGGCTGACCTATGCA	379
E11	F: CAGCATTTGGGCTGTGATGTA R: CGTTCTCTGTTGGAAGGTTGG	400
E12	F: ACCCTGCTCTAGGGAGGTGTC R: CCTCTCCATCCCTTCTACGCT	502
E13	F: AGCCCACTTATCCCCTAGTGC R: ATTTGGGACCTGCTTCATTCA	413

The amplified products were purified with a cycle-pure kit (OMEGA, Bio-Tek, Doraville, GA) and bidirectionally sequenced using an ABI PRISM 3130-xl Genetic Analyzer (Applied Biosystems, Carlsbad, CA) to detect gene mutations. The resulting sequences were compared with the *PAH* gene NCBI reference sequence (NG_008690.1). The variation sites of all patients' sequences were compared to those of the parents, to determine the origin of sequence variability. We identified the biological relationship in de novo mutation pedigrees with Promega PowerPlex 16 HS system (Promega Corporation, Madison, WI, USA).

MLPA analysis

MLPA analysis was performed using the Salsa MLPA Kit P055 PAH (MRC-Holland, Amsterdam, The Netherlands) following DNA sequencing analysis to screen for deletions or duplications in PAH gene exons in seven hyperphenylalaninemics whose genotypes remained partially or completely undetermined. MLPA analysis was carried out by standard protocols. PCR products were identified and quantified using capillary electrophoresis on an ABI 3130-xl Genetic Analyzer with GeneScan software to size the PCR products and to obtain peak areas. For data analysis, the peak sizes and areas were transferred to an Excel file. For normalization, relative probe signals were calculated by dividing each measured peak area by the sum of all peak areas of that sample. The ratio of each relative probe signal from patients compared to control samples was then calculated. An exon deletion was classified as when the ratio was lower than 0.7 and the corresponding samples were subjected to a second MLPA analysis.

STR analysis

The three STR sites PAH-1, PAH-26, and PAH-32 were selected for linkage analysis of prenatal fetuses. The primers for PCR were designed according to the literature (Yao et al., 2007) and the 5' end of the primer was tagged with fluorescein FAM. The PCR reaction was routine. PCR products (5 µL) after degeneration were directly analyzed for fragment analysis on the ABI 3130-xl Genetic Analyzer with the GeneMapper 4.2 software to size the PCR products.

Nomenclature and sequence analysis

The potential mutations were compared with the known disease-causing mutations deposited in disease databases, including the Human Gene Mutation Database (HGMD Professional 2014.3) and the mutation database for the human phenylalanine hydroxylase gene (http://www.pahdb.mcgill.ca/). We excluded single nucleotide polymorphisms (SNPs) by querying 1000 Genomes Data (http://www.1000genomes.org/), the dbSNP database, and the hapmap database. The novel mutations were named according to the Human Genomic Variation Society (HGVS).

Bioinformatic analysis of the mutation sequences

Evolutionary conservation of a non-synonymous variant was estimated with protein sequence alignment generated by ClustalW and compared with that presented by the Ensembl database. The functional consequences of the missense variants were predicted using Polymorphism Phenotyping (PolyPhen-2). Protein Variation Effect Analyzer (PROVEAN) scores, which predict whether a protein sequence variation affects protein function, were also determined.

Prenatal testing

Fetal cells were obtained from transabdominal chorionic villi isolated during the first trimester, by amniocentesis performed during the second trimester, or by umbilical cord blood puncture during the third trimester. In 107 families, *PAH* gene mutation analysis combined with linkage analysis method was used. For the other five families in which the probands were deceased, only *PAH* gene mutation analysis was used.

Maternal contamination from fetal samples and paternity was confirmed by using PowerPlex 16 HS System Kit to ensure the accuracy of linkage analysis, especially for the *de novo* family.

Follow up

Umbilical cord blood was collected at birth for genetic analysis. Heel blood from newborns was screened by Guthrie test. Additionally, urine organic acid analysis using tandem mass spectrometry was also used to confirm results. Developmental assessment of the newborn was performed with age.

RESULTS

Mutations identification and analysis

We identified 61 point mutations or indels and two large deletions in the coding region or adjacent intronic regions of the *PAH* gene on 219 of 224 independent chromosomes. Additionally, we identified five novel mutations (Table 2).

We detected two allele mutations (including homozygotic mutations) in 107 families (including the five families for which the proband is deceased), but five families had only one allele mutation, furthermore, patients with BH4 cofactor deficiency were excluded.

We identified five novel mutations: p.D101N, c.463delC, p.Q172H, p.S250F, and p.L444F.

A substantial proportion of mutant exons and flanking introns (80%) were within exon 7 (32.4%), exon 11 (16.4%), exon 6 (14.6%), exon 12 (9.6%), and exon 3 (8.7%).

We detected the corresponding mutation sites of parents' samples and found one patient carried a de novo mutation; paternity testing identified the mutation as inherited.

Sanger sequencing

The spectrum was composed of 61 different point mutations and indels by Sanger sequencing. The majority was comprised of missense mutations (63.9%), with ten splice-site mutations, six nonsense mutations, and six deletions; the latter three types are classified as null mutations. The 61 identified mutations were distributed across the PAH coding sequence. A substantial proportion of mutant alleles were accounted for by p.R243Q (15.62%), EX6-96A>G (9.82%), p.V399V (7.59%), p.Y356X (6.70%), and p.R413P (5.36%).

We identified a definite diagnosis by detecting two or three mutations on two alleles by using the method in 105 families; the remaining seven families only had one mutant allele.

MLPA analysis

We detected large deletions or duplications of the *PAH* gene by MLPA in seven patients,

Table 2. Spectrum of *PAH* gene mutations detected in the 112 families.

No.	Trivial name (Protein effect)	Systematic name (DNA level)	Location	Characters of mutation	No. of alleles	Relative frequency (%)
	IVS1-3T>C	c.61-3T>C	Intron 1	Splice	1	0.45
	c.131_133delAAG	p.E44_V45delinsV	Exon 2	Deletion	1	0.45
	R53H	c.158G>A	Exon 2	Missense	1	0.45
	IVS2+5G>C	c.168+5G>C	Intron 2	Splice	1	0.45
	S70del	c.208-210delTCT	Exon 3	Deletion	3	1.34
	E79fx13	c.222-225delGAAT	Exon 3	Deletion	3	1.34
	D101N*	c.301G>A	Exon 3	Missense	1	0.45
	H107R	c.320A>G	Exon 3	Missense	2	0.89
	R111X	c.331C>T	Exon 3	Nonsense	10	4.46
	IVS4-1G>A	c.442-1 G>A	Intron 4	Splice	11	4.91
	IVS4+1G>A	c.441+1G>A	Intron 4	Splice	1	0.45
	R155fsX40*	c.463delC	Exon 5	Deletion	1	0.45
	A156P	c.466G>C	Exon 5	Missense	1	0.45
	R158W	c.472C>T	Exon 5	Missense	3	1.34
	F161S	c.482T>C	Exon 5	Missense	3	1.34
	H170Q	c. 510T>A	Exon 5	Missense	2	0.89
	G171R	c. 511G>A	Exon 6	Missense	1	0.45
	Q172H	c. 516 G>T	Exon 6	Missense	1	0.45
	R176X	c.526C>T	Exon 6	Nonsense	6	2.68
	W187X	c.561G>A	Exon 6	Nonsense	1	0.45
	EX6-96A>G	c.611A>G	Exon 6	Splice	22	9.82
	I224T	c.671T>C	Exon 6	Missense	1	0.45
	Q232X	c.694C>T	Exon 7	Nonsense	1	0.45
	R241fsX5	c.722delG	Exon 7	Deletion	1	0.45
	R241C	c.721C>T	Exon 7	Missense	8	3.57
	R241H	c.722G>A	Exon 7	Missense	1	0.45
	L242F	c.724C>T	Exon 7	Missense	1	0.45
	R243Q	c.728G>A	Exon 7	Missense	35	15.62
	G247V	c.740G>T	Exon 7	Missense	5	2.23
	G247R	c.739G>C	Exon 7	Missense	1	0.45
	S250F	c.749C>T	Exon 7	Missense	1	0.45
	R252Q	c.755G>A	Exon 7	Missense	3	1.34
	R252W	c.754C>T	Exon 7	Missense	1	0.45
	G257V	c.770G>T	Exon 7	Missense	1	0.45
	R261Q	c.782G>A	Exon 7	Missense	6	2.68
	F263L	c.787C>T	Exon 7	Missense	1	0.45
	Q267E	c.799C>G	Exon 7	Missense	1	0.45
	P275L	c.824C>T	Exon 7	Missense	2	0.89
	E280K	c.838G>A	Exon 7	Missense	1	0.45
	IVS7+2T>A	c.842+2 T>A	Intron 7	Splice	1	0.45
	IVS8-7A>G	c.913-7A>G	Intron 8	Splice	3	1.34
	A322T	c.964G>A	Exon 9	Missense	1	0.45
	I324N	c.971T>A	Exon 10	Missense	1	0.45
	W326X	c.977G>A	Exon 10	Nonsense	1	0.45
	p.A342>Hfs	c.1024delG	Exon 10	Deletion	1	0.45
	G344S	c.1030G>A	Exon 10	Missense	1	0.45
	A345T	c.1033G>A	Exon 10	Missense	1	0.45
	IVS10-1G>T	c.1066-1 G>T	Intron 10	Splice	1	0.45
	IVS10-11G>A	c.1066-11 G>A	Intron 10	Splice	1	0.45
	Y356X	c.1068C>A	Exon 11	Nonsense	15	6.70
	T380M	c.1139C>T	Exon 11	Missense	1	0.45
	V399V	c.1197A>T	Exon 11	Missense	17	7.59
	R400T	c.1199G>C	Exon 11	Missense	2	0.89
	IVS11+2T>C	c.1199+2 T>C	Intron 11	Splice	1	0.45
	R408W	c.1222C>T	Exon 12	Missense	2	0.89
	R408Q	c.1223G>A	Exon 12	Missense	1	0.45
	R413P	c.1238G>C	Exon 12	Missense	12	5.36
	T418P	c.1252A>C	Exon 12	Missense	1	0.45
	Q419R	c.1256A>C	Exon 12	Missense	1	0.45
	A434D	c.1301C>A	Exon 12	Missense	4	1.79
	L444F*	c.1330 C>T	Exon 13	Missense	1	0.45
	Large deletion		Upstrea +E1	Deletion	1	0.45
			E5	Deletion	1	0.45
	Detected				219	97.77
	Unknown				5	2.23
	Total				224	100

which had only one or no mutation in the *PAH* gene based on the Sanger sequencing. In two of the seven patients, we detected exon deletions. We found a deletion involving a 5' UTR and exon 1 in a single patient and a deletion of exon 5 in another patient (Figure 1).

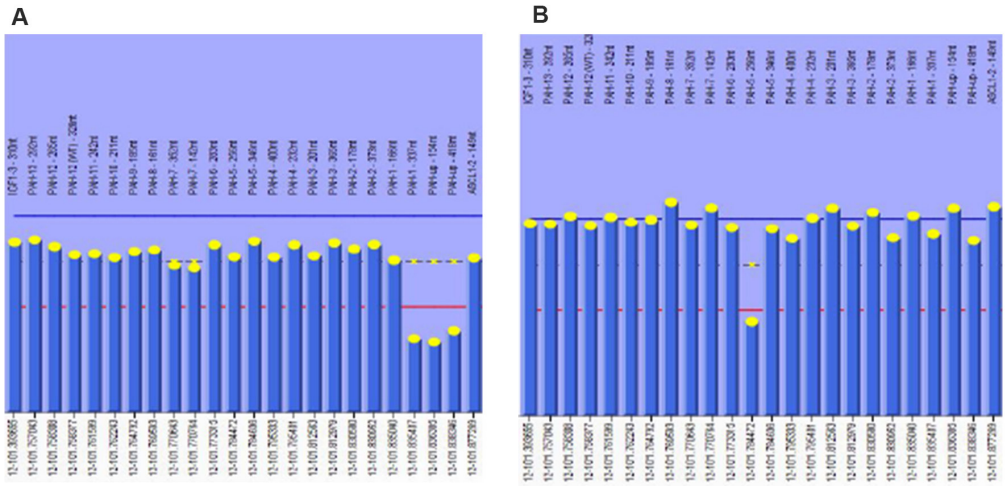

Figure 1. Multiplex ligation-dependent probe amplification analysis. The X-axis of this bar diagram displays specific fragments, and the Y-axis shows normalized ratios. Charts A and B respectively shows heterozygosity for a large deletion of 5'UTR and exon 1, exon 5.

STR analysis

The three STR markers PAH-1, PAH-26, and PAH-32 had distinct heterozygosity, of which PAH-1 detected nine alleles (254, 250, 246, 242, 238, 234, 232, 230, 226), PAH-26 detected six alleles (223, 221, 219, 217, 215, 213), and PAH-32 detected five alleles (148, 146, 144, 142, 138). For 94 families (79.7%), STR analysis could provide effective information; for 19 families, it could only provide exclusive information (16.1%). For the remaining five families, no diagnostic information was provided by STR.

Results of bioinformatic analysis of novel sequence variations

The amino acid sequences of PAH protein in human, chimpanzee, sheep, mouse, and chicken were compared and analyzed. We found that the four novel missense variants sites of the PAH gene are highly conserved in all species (Figure 2). A combination of PROVEAN prediction and PolyPhen-2 showed that the majority of the novel sequence variations may be deleterious mutations rather than polymorphisms (Table 3).

Prenatal diagnosis

For 107 families, the prenatal fetus affected by PKU were detected by PAH gene mutation analysis combined with STR linkage analysis methods. For the remaining five families, we used only *PAH* gene mutation analysis, as the proband is deceased.

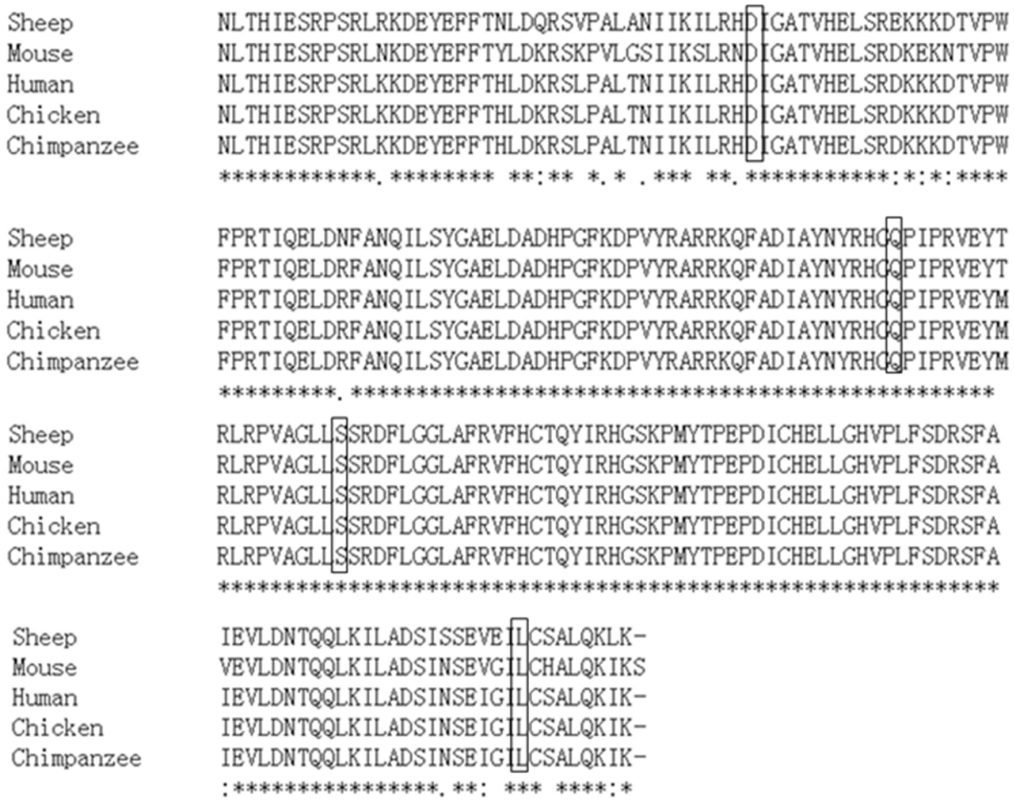

Figure 2. Genetic conservation of the six novel missense mutations. The amino acid sequence of the six codon position D101, Q172, S250, L444 (bounding box) in different species.

Table 3. Novel mutations of *PAH* gene.

NO.	Mutation type	Trivial name (Protein effect)	Systematic name (DNA level)	*In silico* analysis (PolyPhen-2) HumDiv Score	Classification
		D101N	c.301G>A	0.000	Benign
		Q172H	c.516G>T	0.705	Damaging
		S250F	c.749C>T	1.000	Damaging
		L444F	c.1330 C>T	0.999	Damaging

Of the 118 fetuses, 112 were definitively diagnosed by *PAH* gene mutation analysis (diagnostic rate of 91.5%). By STR linkage analysis, 94 of 118 fetuses were definitively diagnosed (diagnostic rate of 79.7%). In addition, the five families carrying one mutant allele were diagnosed using gene mutation analysis and STR linkage analysis.

Among the 118 prenatal genotype fetuses, 31 carried compound heterozygous alleles. Analysis of the aborted tissue confirmed that these fetuses carried two mutant alleles. We identified 58 fetuses as heterozygous, carrying one mutant allele, and 29 fetuses did not carry any detectable mutant allele (Table 4).

Table 4. Prenatal genetic diagnosis in 112 families with PKU.

Families	Proband genotype		Parental genotype		Fetal genotype		Fetal pregnancy outcome
	Mutation 1	Mutation 2	Father source	Mother sources	Mutation 1	Mutation 2	
1	R261Q	R413P	R413P	R261Q	R261Q	R413P	Abortion
2	Y356X		Y356X		Y356X		Normal birth
3	EX6-96A>G	R243Q	EX6-96A>G	R243Q	EX6-96A>G		Normal birth
4	R111X	EX6-96A>G	R111X	EX6-96A>G	R111X		Normal birth
5	IVS4-1G>A	G247V	IVS4-1G>A	G247V			Normal birth
6	E79FX13	R243Q	R243Q	E79FX13	E79FX13		Normal birth
7	EX6-96A>G	R243Q	R243Q	EX6-96A>G	EX6-96A>G		Normal birth
8	R243Q	V399V	R243Q	V399V			Normal birth
9	W326X	Y356X	W326X	Y356X	W326X		Normal birth
10	R252W	Y356X	Y356X	R252W	R252W	Y356X	Abortion
11	EX6-96A>G	IVS10-11G>A	EX6-96A>G	IVS10-11G>A	EX6-96A>G	IVS10-11G>A	Abortion
12	IVS1-3T>C	R243Q	R243Q	IVS1-3T>C	R243Q		Normal birth
13	IVS4-1G>A	Y356X	IVS4-1G>A	Y356X	IVS4-1G>A		Normal birth
14	R261Q	R413P	R413P	R261Q	R261Q	R413P	Abortion
15	R111X	EX6-96A>G	R111X	EX6-96A>G			Normal birth
16	Y356X	A434D	Y356X	A434D	Y356X		Normal birth
17			EX6-96A>G	R243Q	R243Q		Normal birth
18	Y356X	Y356X	Y356X	Y356X	Y356X	Y356X	Abortion
19	G257V	V399V	G257V	V399V	G257V		Normal birth
20	S70del	R243Q	R243Q	S70del	S70del	R243Q	Abortion
21	R111X	R243Q	R243Q	R111X	R111X		Normal birth
22	R243Q	R243Q	R243Q	R243Q			Normal birth
23	R243Q	V399V	V399V	R243Q	R243Q	V399V	Abortion
24	R111X	R413P	R111X	R413P	R413P		Normal birth
25	EX6-96A>G	V399V	EX6-96A>G	V399V	EX6-96A>G		Normal birth
26			R243Q	R243Q	R243Q		Normal birth
27	R243Q	R243Q	R243Q	R243Q			Normal birth
28	R243Q	R261Q	R261Q	R243Q	R243Q	R261Q	Abortion
29	EX6-96A>G	R261Q	R261Q	EX6-96A>G	EX6-96A>G	R261Q	Abortion
30	IVS4-1G>A	R400T	F263L	R400T	F263L	R400T	Abortion
31	IVS4-1G>A	c.722delG	IVS4-1G>A	c.722delG	c.722delG		Normal birth
32	G247V		G247V		G247V		Normal birth
33	IVS4-1G>A	R158W	R158W	IVS4-1G>A	IVS4-1G>A	R158W	Abortion
34	Y356X	R241C	Y356X	R241C	Y356X		Normal birth
35	R243Q	R243Q	R243Q	R243Q	R243Q	R243Q	Abortion
36	R252Q	I324N	I324N	R252Q	R252Q	I324N	Abortion
37	G247V	IVS11+2T>C	G247V	IVS11+2T>C	IVS11+2T>C		Normal birth
38			R243Q				Normal birth
39	EX6-96A>G	R243Q	EX6-96A>G	R243Q	EX6-96A>G	R243Q	Abortion
40	F161S	R243Q	R243Q	F161S	F161S		Normal birth
41	R413P	A434D	R413P	A434D	R413P		Normal birth
42	IVS2+5G>C	G247R	IVS2+5G>C	G247R			Normal birth
43	E79FX13	IVS8-7A>G	IVS8-7A>G	E79FX13	E79FX13		Normal birth
44	EX6-96A>G	V399V	EX6-96A>G	V399V			Normal birth
45	EX6-96A>G	R261Q	R261Q	EX6-96A>G	R261Q		Normal birth
46	R176X	R243Q	R243Q	R176X	R243Q		Normal birth
47	I224T	R243Q	I224T	R243Q			Normal birth
48	IVS4-1G>A	R243Q	IVS4-1G>A	R243Q	R243Q		Normal birth
49	V399V	5' end of exon 1 and the second large fragment LOH	V399V	5' end of exon 1 and the second large fragment LOH	5' end of exon 1 and the second large fragment LOH		Normal birth
50	S70del	Y356X	S70del	Y356X	Y356X		Normal birth
51	R252Q	R413P	R252Q	R413P	R413P		Normal birth
52	F161S	V399V	F161S	V399V	F161S	V399V	Abortion
53	R243Q	R413P	R243Q	R413P	R243Q	R413P	Abortion
54	R243Q	E280K	R243Q	E280K	R243Q	E280K	Abortion
55	IVS4-1G>A	EX6-96A>G	EX6-96A>G	IVS4-1G>A	IVS4-1G>A		Normal birth
56	H170Q	EX6-96A>G	H170Q	P275L	H170Q	P275L	Abortion
57	IVS4-1G>A	P275L	IVS4-1G>A	EX6-96A>G			Normal birth
58	Y356X	A434D	A434D	Y356X			Normal birth
59	EX6-96A>G	R243Q	EX6-96A>G	R243Q	EX6-96A>G		Normal birth
60	A156P	EX6-96A>G	A156P	EX6-96A>G	A156P	EX6-96A>G	Abortion
61			Y356X	R241C	R241C		Normal birth

Continued on next page

Table 4. Continued.

Families	Proband genotype		Parental genotype		Fetal genotype		Fetal pregnancy outcome
	Mutation 1	Mutation 2	Father source	Mother sources	Mutation 1	Mutation 2	
62	EX6-96A>G	R241C	EX6-96A>G	R241C	EX6-96A>G		Normal birth
63	R243Q	G247V	R243Q	G247V	R243Q		Normal birth
64	EX6-96A>G	A345T	EX6-96A>G	A345T			Normal birth
65	V399V	R413P	R413P	V399V	V399V		Normal birth
66	R243Q			R243Q	R243Q		Normal birth
67	EX6-96A>G	R243Q	EX6-96A>G	R243Q	R243Q		Normal birth
68	EX6-96A>G	c.47-48delCT	c.47-48delCT	EX6-96A>G			Normal birth
69	R158W	A434D	A434D	R158W			Normal birth
70	R111X	Y356X	Y356X	R111X	R111X		Normal birth
71	H170Q	P275L	H170Q	P275L	H170Q	P275L	Abortion
72	F263L	R400T	F263L	R400T			Normal birth
73	F161S	V399V	V399V	F161S	F161S	V399V	Abortion
					V399V		Normal birth
74	L242F	V399V	L242F	V399V	L242F		Normal birth
75	Q232X	G247V	G247V	Q232X	Q232X		Normal birth
76			R413P	R243Q	R413P		Normal birth
77	R111X	V399V	V399V	R111X	R111X		Normal birth
78	R243Q	R261Q	R243Q	R261Q	R243Q	R261Q	Abortion
79	R413P	R413P	R413P	R413P	R413P	R413P	Abortion
80	Y356X			Y356X			Normal birth
81	R408W	MLPA E5 LOH	MLPA E5 LOH	R408W	R408W	MLPA E5 LOH	Abortion
82	IVS4+1G>A	EX6-96G>A	IVS4+1G>A	EX6-96G>A	IVS4+1G>A	EX6-96G>A	Abortion
83	Y356X	V399V	Y356X	V399V	Y356X		Normal birth
84	W187X	V399V	V399V	W187X			Normal birth
85	IVS7+2T>A	Y356X	IVS7+2T>A	Y356X	IVS7+2T>A		Normal birth
86	D101N*	Q267E	D101N	Q267E	Q267E		Pregnancy
87	E79FX13	IVS8-7A>G	IVS8-7A>G	E79FX13			Pregnancy
88	R111X	R413P	R111X		R111X		Normal birth
89	R176X	R241C	R241C	R176X			Pregnancy
90	c.463delC*	R243Q	R243Q	c.463delC	R243Q		Pregnancy
91	H107R	R413P	H107R	R413P	R413P		Pregnancy
92	R111X	R176X	R176X	R111X	R176X		Pregnancy
93	H107R	R241C	H107R	R241C			Pregnancy
94	R241H	L444F*	L444F	R241H	R241H		Pregnancy
95	IVS4-1G>A	R176X	IVS4-1G>A	R176X	IVS4-1G>A	R176X	Abortion
96	V399V	V399V	V399V	V399V			Pregnancy
97	R241C	V399V	V399V	R241C	R241C	V399V	Abortion
98	R53H	S250F*	R53H	S250F			Pregnancy
99	Q172H*	R408Q	Q172H	R408Q	R408Q		Pregnancy
100	R252Q	T380M	T380M	R252Q	R252Q		Pregnancy
101	G344S	Q419R	G344S	Q419R			Pregnancy
102	c.131-133delAAG	R408W	c.131-133delAAG	R408W	c.131-133delAAG	R408W	Abortion
103	R176X	A322T	R176X	A322T	A322T		Pregnancy
104	R111X	IVS8-7A>G	IVS8-7A>G	R111X	R111X		Pregnancy
105	R111X	IVS10-1G>T	IVS10-1G>T	R111X			Pregnancy
106	R241C	T418P	R241C	T418P	R241C		Pregnancy
107	S70del			S70del			Pregnancy
108	R243Q	c.1024delG	R243Q	c.1024delG			Pregnancy
109	G171R	EX6-96G>A	G171R	EX6-96G>A			Pregnancy
110	IVS4-1G>A	R158W	IVS4-1G>A	R158W	IVS4-1G>A		Pregnancy
111	R241C	V399V	V399V	R241C	R241C		Pregnancy
112	IVS4-1G>A	R176X	IVS4-1G>A	R176X	IVS4-1G>A	R176X	Abortion
3	EX6-96G>A	R243Q	EX6-96G>A	R243Q	EX6-96G>A	R243Q	Abortion
35	R243Q	R243Q	R243Q	R243Q			Normal birth
58	Y356X	A434D	A434D	Y356X	A434D		Normal birth
65	V399V	R413P	R413P	V399V	V399V	R413P	Abortion
80	Y356X			Y356X			Pregnancy

The fetus in family 88 only carried the mutant allele from the maternal genotype and had a de novo mutation after the paternity was confirmed by using PowerPlex 16 HS.

Family 73 had dizygotic twins identified by STR. Prenatal genotyping showed differences: one of the twins carried compound heterozygous alleles while the other only had one mutant allele.

Following genetic counseling, the guardians chose selective reduction, which was undertook at gestational week 18.

Genetic counseling and follow-up

Following genetic counseling, 30 parents chose to terminate the pregnancy as their fetus carried compound heterozygous mutations. One couple chose to continue the pregnancy and the Phe serum concentration of the neonate was 1650 μM (27.5 mg/dL) by neonatal screening. The 87 carriers with normal fetuses continued their pregnancy, 64 had normal newborns passing neonatal screening, and the remaining 23 are still in pregnancy.

DISCUSSION

The *PAH* gene, located on human chromosome 12q23.2, was cloned by Woo et al. (1983). The gene contains 13 exons and the full-length of encoding region is 90 kb, which encodes a monomer protein that is 452 amino acids. Four PAH proteins interact to form a tetramer, the functional unit of the enzyme, which interacts with four phenylalanine hydroxylase monomers. Previous studies have found over 766 different mutations in the *PAH* gene (HGMD Professional 2014.3). *PAH* gene analysis could provide an important basis for the *PAH* gene spectrum representing different regions and races thereby allowing efficient gene diagnosis, genetic counseling, and prenatal diagnosis of PKU. Additionally, this tool is crucial for further exploration of the relationship between genotype, phenotype, and the function of a gene mutation.

In the study, the highest frequency of allele was exon 7 (33.25%), followed by exons 11, 6, 12, 3, and 5. More than 85% of mutations were detected by screening these six exons, suggesting there are hotspots for mutations in the *PAH* gene in PKU populations of China. Therefore, when screening, priority should be given to the sequencing of exons 6, 7, and 11 of the *PAH* gene, which results in the detection of about 60% mutations. Including exons 3, 5, and 12 in the screening increases the detection to 85% of mutations. This strategy is timesaving, avoids laborious defects of whole coding region sequencing, and saves resources while improving the efficiency of diagnosis.

We identified five novel mutations that have not been indexed in relevant databases. Our major challenge was to distinguish disease-associated mutations from SNPs. The c.463delC frameshift mutation results in the premature termination of *PAH* gene translation and a truncated PAH enzyme. This mutation is generally considered the causative mutation of PKU. To investigate the remaining four missense mutations, we did a conservative analysis between species and then analyzed the mutation function by the PROVEAN prediction and PolyPhen-2 software. The results showed that these missense mutations sequence variation are highly conserved and are inclined to deleterious mutations. The pathogenic properties of the variation depends on the effects of PAH activity in *in vitro* expression. Some of the novel mutations were underlying the *in vitro* expression in our research. Our discovery of novel mutations enriches the human *PAH* gene mutation database.

Sanger sequencing can directly detect point mutations and small indels. MLPA was used to test for the presence of deletions or duplications in the *PAH* gene in a cohort of PKU patients in whom at least one mutation had remained undetected using standard diagnostic tools. After Sanger sequencing and MLPA, only one allele mutation was detected in five patients. We hypothesized that another allele mutation may be present in the regulatory region, exon intron, or distant polyadenylation sites or a SNP may be causing mRNA function changes and influencing the activity of the PAH protein function. Further studies are needed to identify the nature of mutation.

The mutation detection rate was not 100% so prenatal genetic diagnosis was not possible for all families. STR indirect diagnosis can be used in order to complement *PAH* gene sequence analysis and improve the detection rate. Goltsov et al. (1993) found STR polymorphic markers in intron 3, about 700 bp from exon 3, to be associated with the *PAH* gene. The repeating unit ((TCTA)n) represents high polymorphism information and can be used in linkage analysis for prenatal diagnosis. The polymorphism information contents (PIC) PAH-1 locus was between 0.68-0.73 in Chinese populations according to a previous study (Huang et al., 1995), but the single STR site cannot be used in the diagnosis for all families. We selected two additional polymorphic markers, PAH-26 located at telomere 76 kb and PAH-32 located at centromere 85 kb, for linkage analysis, making the detection rate by STR 79.7%. We used multiple fluorescent PCR combined with capillary electrophoresis for STR parting. Multiple PCR allows the amplification of three sites simultaneously, thereby avoiding diagnostic errors caused by sampling error. Additionally, STR parting is simple, accurate by capillary electrophoresis, and improves the working efficiency.

Although combined with multi polymorphic markers, the prenatal diagnosis rate can only achieve 80%. In addition, multiple individuals from the family, especially the proband, are needed for STR linkage analysis, a condition that cannot always be met. Therefore, linkage analysis for prenatal diagnosis of PKU may not always be appropriate.

Hyperphenylalaninemia is a genetically heterogeneous disorder and can be caused by a deficiency of phenylalanine hydroxylase or by a tetrahydrobiopterin deficiency (BH4D). Linkage analysis based on the *PAH* gene does not reliably detect this genetic heterogeneity and can lead to errors. Therefore detection of hyperphenylalaninemia by linkage analysis cannot be performed if hyperphenylalaninemia is not caused by a *PAH* gene mutation.

Of the 118 fetuses from 112 families, we performed prenatal diagnosis through chorionic villus sampling, amniocentesis, or cordocentesis according to the gestational week. Prenatal diagnosis of the fetuses in the 107 core families was performed using *PAH* gene sequencing and STR linkage analysis, resulting in an accurate diagnosis for all fetuses. For the five families in which the proband is deceased, the *PAH* gene of the parents was sequenced and we confirmed that they were carriers. The deceased patients presumably had a *PAH* gene mutation based on the clinical data and *PAH* gene results of the parents. The five fetuses in non-core families were only genotyped by *PAH* gene sequencing. In a single case, a dichorionic- diamniotic twin(DCDA), amniocentesis from two amniotic cavity was performed and genetic diagnosis showed that one fetus was a PKU patient while the other was only a carrier. The guardians chose selective feticide at 18 weeks after informed consent; the other fetus had a normal birth.

For the families in which probands are alive and had differential diagnosis based on BH4 loading test and *PAH* gene sequencing, we recommend that prenatal genetic diagnosis should be performed using *PAH* gene sequencing and STR linkage analysis. While STR linkage analysis can provide prenatal diagnosis information, the accuracy of prenatal diagnosis is improved when combined with gene sequencing. When STR linkage analysis cannot provide information, gene sequencing might allow a prenatal diagnosis. STR analysis and *PAH* gene sequencing complement each other and ensure the accuracy of the prenatal genetic diagnosis. When only one allele of the *PAH* gene has a mutation in the proband, prenatal diagnosis must depend on the linkage analysis. If the proband cannot be tested, parents need to be screened for the *PAH* gene. If the parents are identified as carriers, the prenatal diagnosis can be made by *PAH* gene sequencing. Alternatively, if no mutation is detected in only one or both parents, the prenatal diagnosis cannot be offered.

The recent study of prenatal diagnosis of PKU has been infrequent. The majority of early

studies are based on STR linkage analysis (Fang et al., 1992; Romano et al., 1994; Fan et al., 1999; Kohli et al., 2005; Fazeli et al., 2011), but there have been no cases of complete and large sample research. We investigated 118 cases of prenatal genetically diagnosed fetuses to test *PAH* gene sequencing and STR linkage analysis methods. We confirmed that the methods are accurate and effective for PKU prenatal diagnosis. Furthermore, we propose a new PKU prenatal diagnosis strategy (Figure 3).

PKU is a remediable hereditary disease. Families requiring prenatal diagnosis must be fully informed, provide consent, and should have the doctor objectively explain the prenatal diagnosis results to them; the choice of how to proceed should be determined by the guardians. The prenatal diagnosis fetus should be performed in addition to neonatal screening after birth to verify the accuracy and reliability of the prenatal diagnosis and to allow the opportunity for early treatment if misdiagnosis has occurred.

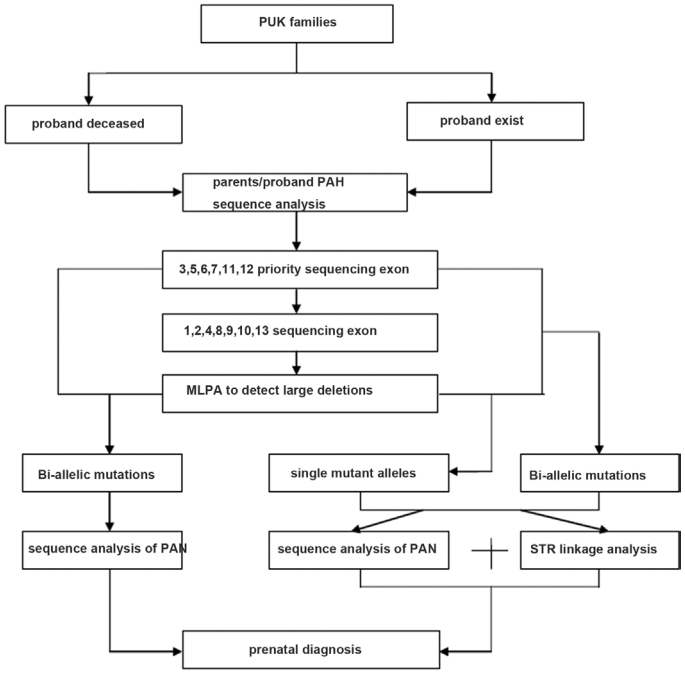

Figure 3. Strategy and flow chart for prenatal diagnosis of PKU.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

We are grateful to the patients and their family members and to all participants in the study for their kind cooperation, especially Professor Zhao DeHua, Neonatal Screening Center, The Third Affiliated Hospital of Zhengzhou University. Research supported by the Henan Province Medical Science and Technique Foundation (Grant #201303015 and #201403021).

REFERENCES

Blau N, van Spronsen FJ and Levy HL (2010). Phenylketonuria. *Lancet* 376: 1417-1427.

Fan GX, Qing LX, Jun Y and Mei Z (1999). Molecular studies and prenatal diagnosis of phenylketonuria in Chinese patients. *Southeast Asian J. Trop. Med. Public Health* 302: 63-65.

Fang B, Yuan L, Wang M, Huang S, et al. (1992). Detection of point mutations of the phenylalanine hydroxylase gene and prenatal diagnosis of phenylketonuria. *Chin. Med. Sci. J*. 7: 205-208.

Fazeli Z and Vallian S (2011). An artifact band frequently associated with variable number of tandem repeat marker at phenylalanine hydroxylase gene: application in carrier detection and prenatal diagnosis of phenylketonuria. *Mol. Biol. Rep*. 38: 3395-3399.

Goltsov AA, Eisensmith RC, Naughton ER, Jin L, et al. (1993). A single polymorphic STR system in the human phenylalanine hydroxylase gene permits rapid prenatal diagnosis and carrier screening for phenylketonuria. *Hum. Mol. Genet*. 2: 577-581.

Gu XF and Wang ZG (2004). Screening for phenylketonuria and congenital hypothyroidism in 5.8 million neonates in China. *Zhonghua Yu Fang Yi Xue Za Zhi* 38: 99-102.

Huang S, Fang B and Chu H (1995). Analysis of short tandem repeats polymorphism in the phenylalanine hydroxylase gene and its application to prenatal gene diagnosis of phenylketonuria. *Zhonghua Yi Xue Za Zhi* 75: 22-24.

Kohli S, Saxena R, Thomas E, Rao P, et al. (2005). Prenatal diagnosis of phenylketonuria. *Indian J. Med. Res*. 122: 400-403.

Romano V, Dianzani I, Ponzone A, Zammarchi E, et al. (1994). Prenatal diagnosis by minisatellite analysis in Italian families with phenylketonuria. *Prenat. Diagn*. 14: 959-962.

Woo SL, Lidsky AS, Güttler F, Chandra T, et al. (1983). Cloned human phenylalanine hydroxylase gene allows prenatal diagnosis and carrier detection of classical phenylketonuria. *Nature* 306: 151-155.

Yao FX, Guo H, Han JJ, Meng Y, et al. (2007). The power of linkage analysis on *PAH* gene in prenatal gene diagnosis is improved with three additional short tandem repeat markers. *Zhonghua Yi Xue Yi Chuan Xue Za Zhi* 24: 382-386.