Clinica Chimica Acta 481 (2018) 132–138

Contents lists available at ScienceDirect

Clinica Chimica Acta

journal homepage: www.elsevier.com/locate/cca

# Mutation spectrum of hyperphenylalaninemia candidate genes and the genotype-phenotype correlation in the Chinese population

Ruifang Wang[a,1], Nan Shen[b,1], Jun Ye[a,1], Lianshu Han[a], Wenjuan Qiu[a], Huiwen Zhang[a], Lili Liang[a], Yu Sun[a], Yanjie Fan[a], Lili Wang[a], Yu Wang[a], Zhuwen Gong[a], Huili Liu[a], Jianguo Wang[a], Hui Yan[a], Nenad Blau[c,**,2], Xuefan Gu[a,*,2], Yongguo Yu[a,*,2]

[a] Department of Pediatric Endocrinology and Genetic Metabolism, Shanghai Institute for Pediatric Research, Xinhua Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai 200092, China
[b] Department of Rehabilitation Medicine, Xinhua Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai 200092, China
[c] Dietmar-Hopp Metabolic Center, University Children's Hospital, 69120 Heidelberg, Germany

ARTICLE INFO

*Keywords:*
Hyperphenylalaninemia
Molecular diagnosis
Mutation spectrum
Genotype-phenotype correlation
Allelic phenotype value
Genotypic phenotype value

ABSTRACT

*Background:* Hyperphenylalaninemia (HPA) is an inherited metabolic disorder that is caused by a deficiency of phenylalanine hydroxylase (PAH) or tetrahydrobiopterin. The prevalence of HPA varies widely around the world.

*Methods:* A spectrum of HPA candidate genes in 1020 Chinese HPA patients was reported. Sanger sequencing, next generation sequencing (NGS), multiplex ligation-dependent probe amplification (MLPA) and quantitative real-time PCR (qRT-PCR) were applied to precisely molecular diagnose HPA patients. The allelic phenotype values (APV) and genotypic phenotype values (GPV) were calculated in PAH-deficient patients based on a recently developed formula.

*Results:* Apart from genetic diagnoses confirmed in 915 HPA patients (89.7%) by Sanger sequencing, pathogenic variants were discovered in another 57 patients (5.6%) through deep detections (NGS, MLPA and qRT-PCR). We identified 196, 42, 10 and 2 variants in *PAH*, *PTS*, *QDPR* and *GCH1*, respectively. And a total of 47 novel variants were found in these genes. Through the APV and GPV calculations, it was found that the new GPV system was well correlated with metabolic phenotypes in most PAH-deficient patients.

*Conclusions:* More HPA candidate variants were identified using new molecular diagnostic methods. The new APV and GPV system is likely to be highly beneficial for predicting clinical phenotypes for PAH-deficient patients.

## 1. Introduction

Hyperphenylalaninemia (HPA) is the most common inborn error of amino acid metabolism all over the world. The occurrence of HPA varies widely around the world. The average incidence in Europe is about 1:8000 [1], and in the USA is 1:15,000 [2]. While in mainland China the reported prevalence of this disease is approximately 1:12,000 [3,4]. HPA is mainly caused by a deficiency of hepatic phenylalanine hydroxylase (PAHD, MIM # 261600), which catalyzes the conversion of L-phenylalanine (Phe) to L-tyrosine (Tyr). The cofactor tetrahydrobiopterin (BH4) and molecular oxygen are required for this enzymatic process [5,6]. Guanosine triphosphate cyclohydrolase I (GTPCH, MIM# 233910), 6-pyruvoyl-tetrahydropterin synthase (PTPS, MIM# 261640), sepiapterin reductase (SR, MIM# 612716), pterin-4-alpha-carbinolamine dehydratase (PCD, MIM# 264070) and dihydropteridine reductase (DHPR, MIM# 261630) are essential for the

*Abbreviations:* HPA, hyperphenylalaninemia; PAH, phenylalanine hydroxylase; BH4, tetrahydrobiopterin; NGS, next generation sequencing; MLPA, multiplex ligation-dependent probe amplification; qRT-PCR, quantitative real-time PCR; APV, allelic phenotype values; GPV, genotypic phenotype values; PAHD, deficiency of hepatic phenylalanine hydroxylase; Phe, L-phenylalanine; Tyr, L-tyrosine; GTPCH, guanosine triphosphate cyclohydrolase I; PTPS, 6-pyruvoyl-tetrahydropterin synthase; SR, sepiapterin reductase; PCD, pterin-4-alpha-carbinolamine dehydratase; DHPR, dihydropteridine reductase; MHP, mild HPA; MPKU, mild phenylketonuria; CPKU, classic phenylketonuria; AV, assigned values; DBS, dried blood spot; ExAC, Exome Aggregation Consortium

* Correspondence to: X. Gu and Y. Yu, Room 801 Sci&EduBldg, 1665 Kongjiang Rd, Yangpu District, Shanghai 200092, China.
** Correspondence to: N. Blau, Dietmar-Hopp Metabolic Center, University Children's Hospital, Im Neuenheimer Feld 669, D-69120 Heidelberg, Germany.
*E-mail addresses:* nenad.blau@med.uni-heidelberg.de (N. Blau), guxuefan@xinhuamed.com.cn (X. Gu), yuyongguo@shsmu.edu.cn (Y. Yu).
[1] The first three authors contributed equally to this work, and should be considered as co-first author.
[2] The last three authors contributed equally to this work.

https://doi.org/10.1016/j.cca.2018.02.035
Received 16 November 2017; Received in revised form 10 February 2018; Accepted 27 February 2018
Available online 28 February 2018

biosynthesis and regeneration of BH4. Classical forms of BH4 deficiency, caused by mutations in the genes encoding GTPCH, PTPS, PCD and DHPR, can lead to HPA [7,8]. However, in contrast to these forms of BH4 deficiency, SR deficiency may present a phenotype without HPA [9]. All infants should undergo neonatal screening programs in the first few days after birth to detect HPA. Blood Phe concentrations below 120 μmol/L denote normal. If the Phe levels increase to 120–600 μmol/L, patients are diagnosed with mild HPA (MHP). Those patients with blood Phe concentrations of 600–1200 μmol/L are considered mild phenylketonuria (MPKU), and Phe concentrations above 1200 μmol/L indicate classic phenylketonuria (CPKU) [5,10].

Currently, a combination of different sequencing methods, such as conventional Sanger sequencing and next generation sequencing (NGS) technology, allow physicians to obtain patients' exact genotypes in a short time [11–13]. The genotypes related to HPA, especially PAH genotyping, can be useful for the establishment of a genotype-phenotype correlation and the application of precise therapy for these patients [14]. A correlation between PAH genotypes and metabolic phenotypes has been reported for some *PAH* mutations [15–17]. Thus, many previous studies have attempted to develop a system that can predict metabolic phenotypes from particular alleles [18,19]. The ability to predict phenotypes contributes to the better characterization of PKU patients. The Guldberg's arbitrary assigned values (AV) system is based on scoring of gene variants and is widely used for the phenotype prediction of *PAH* variants. It classifies HPA as different metabolic phenotypes using four levels (AV1 = CPKU, AV2 = moderate PKU, AV4 = MPKU and AV8 = MHP) [18]. However, in the Guldberg et al.'s study, researchers enrolled 686 HPA patients from seven European centers. The different methods and criteria used for diagnosis and classification of HPA in different medical centers may have caused to the inconsistencies in the genotypes and phenotypes of some variants.

In the present study, a spectrum of gene mutations related to HPA was reported from a comprehensive survey of 1020 HPA patients in mainland China. Sanger sequencing was firstly performed to find causative gene variants in these patients. Then, NGS technology, multiplex ligation-dependent probe amplification (MLPA) and quantitative real-time PCR (qRT-PCR) were used to make further precise molecular diagnoses. Among all 1020 HPA patients, 808 had a PAH deficiency that caused HPA. For patients with *PAH* variants, we applied a recently developed linear allelic phenotype values (APV) and genotypic phenotype values (GPV) system to predict PAH genotype-phenotype relationship in Chinese PKU population.

## 2. Materials and methods

### 2.1. Patients and phenotypes

A total of 1020 nonconsanguineous Chinese HPA patients (554 males and 466 females) in the pediatric endocrinology clinic of Shanghai Xinhua Hospital were investigated from 2011 to 2015. These patients were identified through a neonatal screening program. The plasma Phe concentrations were measured by tandem mass spectrometry from dried blood spot (DBS) samples before starting treatment. All patients with the maximum pretreatment blood Phe concentrations above 120 μmol/L were enrolled. In addition, urinary pterin analysis using high performance liquid chromatography and a DHPR activity assay [20] on DBS samples were also carried out to distinguish between PAH deficiency and tetrahydrobiopterin deficiency. Informed consents were obtained from all participants or their parents in prior to the commencement of the study. And this research was prospectively reviewed and approved by Ethics Committee of Xinhua Hospital affiliated to Shanghai Jiaotong University School of Medicine (Approval Number XHEC-D-2017-055).

### 2.2. Genome sequencing

Sanger sequencing, NGS, MLPA and qRT-PCR were performed to detect *PAH* (NM_000277.1), *PTS* (NM_000317.2), *QDPR* (NM_000320.2), *GCH1* (NM_000161.2) and *PCBD1* (NM_000281.3) gene, which encode PAH, PTPS, DHPR, GTPCH and PCD, respectively. These methods for detecting mutations are summarized below briefly.

For Sanger sequencing, genomic DNA was extracted from the peripheral blood mononuclear cells of patients by a density centrifugation (Ficoll-Paque) method. The primers used for Sanger sequencing were designed for the PCR amplification of all exons and exon-intron boundaries of the five genes mentioned above using Primer Premier 5.0 software (Premier Biosoft International, Palo Alto, CA, US). The PCR products were purified and sequenced bidirectionally using an ABI 3730xl DNA analyzer (Applied Biosystems, Foster City, CA, USA).

NGS was used for patients whose genetic diagnosis of HPA could not be confirmed by Sanger sequencing. A Ion AmpliSeq™ Custom Primer Panel that contained 52 amplicons was designed and ordered. This custom panel can targeted all of the exons (padding: ± 10 bp) of *PAH*, *PTS*, *GCH1*, *PCBD1* and *QDPR* gene, with the exception of 66 bp in exon 1 of *GCH1* gene. To construct the libraries, 10–15 ng of each DNA sample was used with an Ion AmpliSeq™ Library Kit v2.0 (Cat. 4,475,345, Life Technologies, CA, USA). Gene mutations were sequenced by the Ion Personal Genome Machine® System (Life Technologies, CA, USA).

After sequencing all samples, the data were analyzed on the Ion Torrent server using Torrent Suite v4.2.1 (Life Technologies, CA, USA). First, each sample read was classified according to the special barcode, signal processing, basic calling, read filter and sequence alignment. Second, the generated VCF files for each sample were annotated with the Ion reporter software (https://ionreporter.thermofisher.com). Finally, the annotated results were further analyzed by comparing them with dbSNP (https://www.ncbi.nlm.nih.gov/projects/SNP), 1000 Genomes (http://www.internationalgenome.org/), Exome Aggregation Consortium (ExAC, http://exac.broadinstitute.org/), PAHvdb (http://www.biopku.org/home/pah.asp) and the PNDdb database (http://www.biopku.org/home/pnddb.asp).

The SALSA MLPA kit P055 PAH (Cat. P055-100R, MRC Holland, Amsterdam, Netherlands) was used to detect large deletions or duplications in patients with only one variant site of *PAH* or in those with no mutations found. The assay operation was performed according to the manufacturer's protocol. After DNA denaturation, ligation of the probes and multiplex PCR amplification, the PCR products were detected by the ABI 3130 Genetic Analyzer (Applied Biosystems, Foster City, CA, US) using capillary electrophoresis. The raw data were analyzed by Coffalyser data analysis software (MRC Holland, Amsterdam, Netherlands). DNA samples from four healthy individuals were used as normal control samples in each run.

Possible deletions in patients with only one mutation site of *PTS* gene were detected with qRT-PCR. Seven pairs of primers were designed for qRT-PCR using Primer 3.0 software (http://biotools.umassmed.edu/bioapps/primer3_www.cgi) (Supplementary Table 1). Detecting operation was performed on the ABI 7500 real-time PCR system (Applied Biosystems, Foster City, CA, US).

### 2.3. Pathologic analysis of novel mutations of PAH and PTS genes

For novel variants found in our patients, the evolutionary conservation was estimated with Clustal X and the effects of amino acid substitutions were evaluated using three online tools: SIFT (http://sift.jcvi.org), PolyPhen (http://genetics.bwh.harvard.edu/pph/) and Mutation Taster (http://www.mutationtaster.org/).