Gene 529 (2013) 80–87

Contents lists available at ScienceDirect

Gene

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

Short Communication

Variations in genotype–phenotype correlations in phenylalanine hydroxylase deficiency in Chinese Han population☆,☆☆

Tianwen Zhu [a], Jun Ye [b], Lianshu Han [b], Wenjuan Qiu [b], Huiwen Zhang [b], Lili Liang [b], Xuefan Gu [b,*]

[a] Department of Neonatal Medicine, Xin-Hua Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, China
[b] Department of Endocrinology and Genetic Metabolism, Xin-Hua Hospital, Shanghai Institute of Pediatric Research Affiliated to Shanghai Jiao Tong University School of Medicine, China

ARTICLE INFO

Article history:
Accepted 22 July 2013
Available online 7 August 2013

Keywords:
PAHD
Chinese Han population
Phenotype–genotype correlation
Complex phenotypes

ABSTRACT

Background: The value of genotyping to predict variant phenotypes in patients with phenylalanine hydroxylase (*Pah*) deficiency is a matter of debate. However, there exists no comprehensive population relationship study focused on the Han Chinese.
Methods: We analyzed genotype–phenotype correlation for 186 different genotypes in 338 unrelated Chinese patients harboring 109 different *Pah* mutations. Two systems were used in this process. The first was a phenotype prediction system based on arbitrary values (AV) attributed to each mutation. The second was a pair-wise correlation analysis. The observed phenotype for AV analysis was the corresponding metabolic phenotype stratified according to the pretreatment phenylalanine (Phe) value.
Results: We found that the observed phenotype matched the predicted phenotype in 54.41% of 272 patients for whom AV information was available; the highest degree of concordance (61.83%) was found in patients with null/null genotypes, whereas the lowest "concordance rate" (32.69%) was observed for patients with expected mild-PKU phenotype. There are repeated inconsistencies for such mutations as R241C, R243Q, R261Q, V388M, V399V, R408Q, A434D and EX6-96A>G which are associated with variable phenotypes in patients with identical genotype. Significant correlations were disclosed between pretreatment Phe values and predicted residual activity ($r = -0.45643$, $P < 0.0001$) or AV sum ($r = -0.59523$, $P < 0.0001$).
Conclusion: Our study supports the notion that the *Pah* mutation genotype is the main determinant of metabolic phenotype in most patients in a particular population, and provided novel insights into the values that underpin the subsequent treatment and the prognosis of PKU in Chinese.

1. Introduction

Phenylalanine hydroxylase (*Pah* [MIM 261600]) deficiency (PAHD) is an autosomal recessive disorder that results in intolerance to the dietary intake of phenylalanine (Mitchell et al., 2011). It is the most common inborn error of amino acid metabolism in Chinese, with an incidence of about 1:11,572 (Zhan et al., 2009), which is similar to that in Caucasian populations (Zschocke, 2003). The key metabolic feature of this disease is elevated serum concentrations of phenylalanine, leading to the mental retardation with varying degrees if left untreated.

So far, more than 500 different mutant alleles, which cause different levels of reduction in the catalytic activity of the enzyme, have been identified at the *Pah* locus (http://www.pahdb.mcgill.ca/) with wide variation of their frequency and genotypic distribution, generating a wide spectrum of phenotypes ranging in severity from classic phenylketonuria (PKU) to variant PKU or mild hyperphenylalaninemia (MHP). There has been a hope that delineation of genotypes would enable the prediction of variant phenotypes in the case of human genetic disease, which would have added value for prognosis and treatment. Nonetheless, reports on genotype–phenotype relationship in PAHD in some European (Daniele et al., 2007; Groselj et al., 2012; Mallolas et al., 1999) and Oriental populations (Chien et al., 2004; Okano et al., 2011; Qu et al., 2008) often showed no robust correlation due to the high allelic heterogeneity and broad phenotypic variability.

Recently, a more formalized system by Guldberg et al. (1998) was developed for estimating genotype–phenotype correlations based on the prediction of the phenotypic impact of each mutation and was adopted in some studies. For some populations, such as Brazilian and

Abbreviations: AV, arbitrary value; MHP, mild hyperphenylalaninemia; *Pah*, phenylalanine hydroxylase; PAHD, phenylalanine hydroxylase deficiency; Phe, phenylalanine; PKU, phenylketonuria; PRA, predicted residual PAH activity.

☆ Funding source: This work was supported by grants from the Major Program of Shanghai Committee of Science and Technology (11dz195030), the Shanghai City Health Bureau project (20124104), the National Natural Science Foundation of China (81070700, 81200654), the National Key Technology R&D Program(2012BAI09B04) and the Shanghai Jiao Tong University School of Medicine Fund (11XJ22002).
☆☆ Financial disclosure: All authors have no financial relationships relevant to this article to disclose.

* Corresponding author at: Department of Endocrinology and Genetic Metabolism, XinHua Hospital, Shanghai Institute for Pediatric Research Affiliated to Shanghai Jiao Tong University School of Medicine, Kongjiang Road 1665#, Shanghai 200092, China. Tel.: +86 21 65011012; fax: +86 21 65791316.
E-mail address: gu_xf53@yahoo.com.cn (X. Gu).

0378-1119/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.gene.2013.07.079

Lithuanian, in which some particular allele frequencies were highly skewed (Acosta et al., 2001; Kasnauskiene et al., 2003), the system proved to be highly useful, indicating that the efficiency of the method might vary depending on the set of mutations in a specific population.

In our previous study of the Chinese Han population (Zhu et al., 2010), 79 different mutations were identified in 212 unrelated patients with 8 mutations accounting for two-thirds of the identified ones, which facilitated investigation of their phenotypic effect. The present study was sought to examine genotype–phenotype correlation for 186 different genotypes in 338 unrelated Chinese patients (the genetic analysis for some of them had been made in our previous study).

2. Materials and methods

2.1. Patients and phenotypic classification

DNA samples were collected in the Department of Pediatric Endocrinology and Genetic Metabolism, Xin-Hua Hospital Affiliated to Shanghai Jiao Tong University School of Medicine. This division treats and provides follow-up for more than two thousand PKU patients with known ethnic origin. 338 unrelated Chinese patients with two presumably causative alterations in the *Pah* gene from January 2006 through December 2012 were recorded in our database. In all patients, hyperphenylalaninemia (HPA) had been detected by either national screening (57.10%, 193/338) or presence of neurological deterioration at elder age (42.90%, 145/338) with a plasma phenylalanine (Phe) cut-off level of 120 μmol/L. A defect in the synthesis or recycling of tetrahydrobiopterin was excluded by analysis of urinary pterins and dihydropteridine reductase activity in erythrocytes.

Metabolic phenotype for each patient was stratified according to the plasma Phe concentration before treatment, and all of which applied in this study were the maximum pretreatment values. Patients were classified as classic PKU (Phe, more than 1200 μmol/L), moderate PKU (Phe, 900 to 1200 μmol/L), mild PKU (Phe, 600 to 900 μmol/L) and MHP who keeps their Phe levels below 600 μmol/L on a free diet.

The local Ethics Committee approved this study and informed consents were obtained from the parents of these patients enrolled.

2.2. Genotype analysis and mutation nomenclature

Genomic DNA was isolated from peripheral blood samples, 13 exons and related intronic boundaries of *Pah* gene were amplified, and all PCR products were scanned for mutations by direct sequence analysis.

Mutations were referred to by their "trivial names", as registered in the *Pah* Mutation Analysis Consortium Database. Genotypes were coded in two different ways. Initially, a simple sequential order was used. Thereafter, codes based on the residual activity of each allele were constructed. In this study, the nonsense, frame shift, splicing and those missense mutations with an enzyme activity in vitro less than 3% were given a definition of null *Pah* activity which were grouped in Table 1. Alleles classified as null or missense with some residual activity, were distributed among three genotype categories: null/null, null/missense (functionally heterozygous) and missense/missense in the order of increasing predicted residual *Pah* activity (PRA).

2.3. Data on in vitro expression analysis

PRA was recorded from data provided from in vitro experiments using recombinant expressed mutant proteins in eukaryotic cells. Expression data were compiled mainly from the *Pah* Mutation Analysis Consortium Database (http://www.pahdb.mcgill.ca/) or other published papers (Gersting et al., 2008; Gjetting et al., 2001; Guldberg et al., 1998; Kim et al., 2006; Okano et al., 1998; Pey et al., 2003; Waters et al., 1998). Mean PRA values, the average of the sum of activities of both alleles, were calculated for each genotype. The mutations present in 66 individuals lack expression data. These patients were excluded from the analyses that depended on in vitro expression information.

Table 1
Classification, by type, of 53 null mutations.

Classification of 41 null mutations by type	
Mutation type	Mutations
Missense (n = 16)	R413P[a], E280K[a], L255S[a], R252Q[a], R270S[a], T278I[a], R408W[b], R252Wb, A259T[b], G247V[b], R243Q[b], A342T[b], R158Q[b], A156P[c], L385P[c], R400K[c]
Protein truncation (n = 15)	R111X, Y356X, Y166X, W326X, R176X, R261X, C375X, E228X, R243X, Q232X, W187X, S411X, Q172X, Q301X, Y414X
Deletion (n = 6)	S70del, 190-194delCACAT, F39del, 1024delG, 541-543delGAG, 540delGGAGG
Frameshift (n = 1)	R241>Pfs
Splice defective (n = 15)	EX6-96A>G, V399V, IVS4-1G>A, IVS6-1G>A, VS7+2T>A, IVS4+3G>C, IVS5+1G>A, IVS4+2T>A, IVS10-14C>G, IVS2+5G>C, IVS7+1G>A, IVS7+5G>A, IVS12+6T>A, IVS12+4A>G, IVS5-2A>G

[a] A null phenotype effect was declared when enzyme activity was demonstrated to be below the level of detection in the in vitro system (typically, <3% or <1% of normal.
[b] From data in literature, although also represented in other categories at lower frequencies.
[c] Identified in at least one patient with classic PKU in the homoallelic state and therefore, may be formally classified as a null mutation.

2.4. Phenotypic prediction system

Mutations were assigned to one of the four-phenotype categories (classic, moderate, mild, and MHP), according to Guldberg et al. (1998). An arbitrary value (AV) was assigned to each mutation: AV = 1 for classical PKU mutation; AV = 2 for moderate PKU mutation; AV = 4 for mild PKU mutation, and AV = 8 for MHP mutation. Phenotypes resulting from a combination of the two mutant alleles were expressed as the sum of the two mutations' AVs. Sixty-six individuals, whose mutations without AV estimates derived from our data (Appendix A) or the literature, were excluded from those analyses that depended on AV information.

2.5. Statistical analysis

Statistical analysis was implemented with SPSS 13.0®. Spearman correlation was estimated between the pretreatment Phe levels and mean PRA, AV sum respectively. A significance level of $P < 0.05$ was considered for all the analyses. The differences of pretreatment Phe concentrations among genotype groups based on PRA were evaluated by one-way ANOVA.

3. Results

3.1. Mutation and phenotype distribution

A total of 338 unrelated patients were investigated. 109 different mutations were discovered, including 72 (66.06%) missense mutations, 15 nonsense, 15 splice-site and 7 frame-shift deletions. The eight most prevalent mutations of the 676 alleles were R243Q representing 24.11%, EX6-96A>G with 10.65%, R241C with 7.54%, R111X with 5.47%, IVS4-1G>A with 5.33%, V399V with 4.88%, Y356X with 4.88% and R413P with 4.44%. The first two accounted for one third of the identified mutations, and the next six for another third. The defined mutations were distributed from exon 2 through 7, and in exons 11 and 12. On the basis of individual data on pretreatment Phe levels, the patients were assigned to the four arbitrary metabolic phenotype categories, with 143 (42.31%) as classic PKU, 80 (23.67%) as moderate PKU, 71 (21.01%) as mild PKU and 44 (13.02%) as MHP.