Metabolic Brain Disease (2019) 34:1547–1555
https://doi.org/10.1007/s11011-019-00461-w

ORIGINAL ARTICLE

The study of the full spectrum of variants leading to hyperphenylalaninemia have revealed 10 new variants in the *PAH* gene

I. Kuznetcova[1] · P. Gundorova[1] · O. Ryzhkova[1] · A. Polyakov[1]

Received: 21 February 2019 / Accepted: 8 July 2019 / Published online: 22 July 2019

Abstract

This study presents further research into the spectrum of variants in genes responsible for the development of phenylketonuria (PKU) and hyperphenylalaninemia (HPA) in patients in Russia. After a study of 25 frequent variants, 293 patients (327 chromosomes without detected variants) from among 1265 probands still had no confirmed diagnosis. A study involving methods of next generation sequencing (NGS) of *PAH*, *PTS*, *GCH1*, *PCBD1*, *QDPR*, *SPR* and *DNAJC12* genes to search for point mutations and multiplex ligation-dependent probe amplification (MLPA) methods to search for gross deletions were conducted for these patients. Among 327 chromosomes without identified variants, variants in the *PAH* gene were found on 260 chromosomes, and variants in the *PTS* gene were found on 10 chromosomes. On 10 chromosomes gross deletions by the MLPA method were detected. 104 rare variants of the *PAH* gene, including 10 variants not previously described, and 6 variants of the *PTS* gene were revealed. The NGS method revealed additional *PAH* gene variants on 10.3% of chromosomes and *PTS* gene variants on 0.4%. Gross deletions of the *PAH* gene were revealed in 0.5% of chromosomes. Thus, the most complete understanding of the spectrum of variants leading to the development of the PKU and HPA in Russia with the use of all methods available today has been obtained. Such a detailed study of the spectrum of rare variants on the genetic material from Russia was undertaken for the first time.

Keywords Next generation sequencing · Tetrahydrobiopterin · Phenylketonuria · Hyperphenylalaninemia

Introduction

Hyperphenylalaninemia (HPA) is a group of hereditary heterogeneous autosomal recessive disorders characterized by an increase in the level of phenylalanine (Phe) in blood. The cause of the increase of Phe level in blood is a disorder of the phenylalanine hydroxylase (PAH) enzyme responsible for the metabolism of phenylalanine. Hyperphenylalaninemias are classified according to the molecular-genetic cause of disease. Phenylketonuria (PKU), tetrahydrobiopterin-deficient (BH4-deficient) HPA and non-BH4-deficient HPA are distinguished. Phenylketonuria (OMIM #261600) is the most common type of HPA (about 98%), caused by phenylalanine hydroxylase (PAH) dysfunction due to variants in the *PAH* gene (Zurfluh et al. 2008). BH4-deficient HPA (HPABH4) is caused by variants in genes of proteins responsible for synthesis and metabolism of BH4. There are 5 types of BH4-deficient HPA classified. HPABH4A (OMIM #261640) is caused by variants in the *PTS* gene, which encodes 6-pyruvoyltetrahydrobiopterin synthetase enzyme (Dudesek et al. 2001). The causes of occurrence of HPABH4B (OMIM #233910) are variants of the *GCH1* gene, which encodes GTP-cyclohydrolase 1 (Dhondt et al. 1985). HPABH4C (OMIM #261630) occurs due to variants in the *QDPR* gene, which encodes dihydropteridine reductase (Grobe et al. 1978). Variants in the *PCBD1* gene, which encodes pterin-4α-carbinol-amine-dehydratase, lead to the occurrence of HPABH4D (OMIM #264070) (Thony et al. 1998). In a group of BH4-deficient disorders there is also a DOPA-responsive dystonia due to the deficiency of sepiapterin reductase (OMIM #612716) encoded by the *SPR*

✉ I. Kuznetcova
Iradris@yandex.ru

[1] Federal State Funded Research Institution "Research Centre of Medical Genetics", Moscow, Russian Federation

gene which results in a severe dopamine and serotonin deficiencies in the central nervous system caused by a defect in BH4 synthesis (Bonafe et al. 2001). The connection of hyperphenylalaninemia with variants in the *DNAJC12* gene was detected in 2017. This condition is called non-BH4-deficient HPA (OMIM #617384). Since DNAJC12 and BH4 proteins both play role of chaperones of the PAH enzyme, clinical features of the DNAJC12 deficiency are similar to the group of BH4-deficient disorders (Anikster et al. 2017). One of the objectives of the study is to determine the percentage of rare forms of HPA in Russia, which arise from the disturbance of BH4 synthesis and makes 1–2% of all the HPA globally.

Since HPA leads to irreversible changes at an early age of the patient, therapy shall be assigned immediately. The most effective PKU treatment has been a low phenylalanine diet. BH4-deficient HPAs are treated with sapropterin (pharmacological analogue of tetrahydropterin) in combination with levodopa drugs (van Wegberg et al. 2017). It is noted that BH4 therapy is successful in patients with BH4-sensitive PKU. To determine the sensitivity to sapropterin, loading tests with the drug are conducted. Currently, loading tests are not a routine procedure in Russia. The potential response to treatment can also be predicted based on the information about *PAH* gene variants found in the patient. Variants with the residual activity of the PAH enzyme exceeding 10% is referred to as "mild" variants, in which the enzyme activity below 10% is referred to as "severe" ones. If a patient has two "severe" variants in his/her genotype, loading test will be ineffective. Such patients do not respond to the treatment by tetrahydrobiopterin analogues. Loading tests are recommended if at least one "mild" variant is found in the patient's genotype (van Wegberg et al. 2017).

In Russia, in the DNA-diagnostics laboratory of Federal State Funded Research Institution "Research Centre of Medical Genetics", the program of genotyping of patients with PKU and HPA is being implemented. From December 2016 to January 2018, the material of 1265 unrelated probands was studied. The search for 25 frequent variants in the *PAH* gene by the allele-specific MLPA method was carried out for all samples.

Currently, HPA diagnostics is carried out in several stages. The first stage involves the search for 25 frequent variants in the *PAH* gene (p.Ser16* (c.47_48delCT), p.Leu48Ser, IVS2+5G>A, IVS2+5G>C, p.Arg111*, IVS4+5G>T, EX5del4154ins268, p.Arg158Gln, p.Asp222* (c.664_665delGA), p.Arg243Gln, p.Arg243*, p.Arg252Trp, p.Arg261Gln, p.Arg261*, p.Glu280Lys, p.Pro281Leu, p.Ala300Ser, p.Ile306Val, p.Ser349Pro, IVS10-11G>A, p.Glu390Gly, p.Ala403Val, p.Arg408Trp, p.Tyr414Cys, IVS12+1G>A), the total allelic frequency of which is 86.1% (Gundorova et al. 2019). If the diagnosis is not confirmed, Sanger sequencing of *PAH* and HPABH4 genes can be performed. This method is time consuming and resource intensive. In order to shorten the time of analysis and processing of the result, application of the next generation sequencing method is recommended. Unlike the Sanger sequencing method, this method allows analyzing a large number of genes in one reaction. In addition, the above method will help to optimize the existing HPA diagnostics protocol in Russia.

Materials and methods

Patient selection

DNA samples of 1265 unrelated probands with "phenylketonuria" or "hyperphenylalaninemia" diagnoses were studied. Probands with these diagnoses were directed by the treating physicians from the regional medical genetic consulting centers based on the results of biochemical screening and evaluation of the clinical findings. The diagnosis were confirmed for probands using biochemical analysis in accordance with the European guidelines on phenylketonuria (van Wegberg et al. 2017). These patients were included in the Russian phenylketonuria neonatal screening and treatment audit. According to the PKU patients genotyping program, all probands were analyzed for the presence of 25 frequent *PAH* gene variants. Ethics Committee of the Federal State Budgetary Institution "Research Centre for Medical Genetics" had approved the study with the Protocol No. 2/2 at the meeting on February 5, 2018. Informed consent was obtained from all individual participants included in the study.

According to the program results, 293 patients had not their diagnosis confirmed. These patients were divided into two groups. The first group consisted of 259 patients with one variant detected in the course of search for the 25 most frequent *PAH* gene variants. The second group consisted of 34 patients without variants detected. Thus, the analysis included 327 chromosomes from patients without variants detected.

Next generation sequencing

For the search of rare variants, the Next Generation Sequencing user panel Ampliseq™, covering the DNA coding sequence, exon boundaries, and also partially covering 5′ and 3′ untranslated regions of *PAH*, *PTS*, *GCH1*, *PCBD1*, *QDPR*, *SPR* and *DNAJC12* genes, was developed using Ion AmpliSeq™ software. The panel consisted of two separate PCR pools of primers with a total number of 68. The average length of the amplicons of the panel obtained was 158 bps. For each sample, a library was created using the AmpliSeq™ Library Kit 2.0 commercial kit (Life Technologies, USA) and following the manufacturer's instructions. The samples were marked with unique barcodes (Ion Xpress™ Barcode Adapters Kit, Life Technologies, USA), and then united in equimolar concentrations. Libraries were prepared for

sequencing in automatic mode on the Ion Chef™ device (Termofisher Scientific, USA). Next generation sequencing was carried out on the Ion S5™ device (Termofisher Scientific, USA). Sequencing data processing was performed using the standard automated algorithm offered by Termofisher Scientific (Torrent Suite™, USA) and GeneTalk software. The overall coverage of genes amounted to: *PAH*-100%, *PTS*-98%, *GCH1*–87,2%, *PCBD1*–94%, *QDPR*-100%, *SPR*-82,3%, *DNAJC12*–100%. Uncovered regions of the *PTS* gene were additionally analyzed by the Sanger direct automatic sequencing method. The clinical significance (pathogenesis) of the variants revealed was evaluated on the basis of recommendations for the interpretation of data received by next generation sequencing methods (Richards et al. 2015). The variants revealed were analyzed by reference sequences NM_000277.2 and NM_000317.2 of the *PAH* and *PTS* genes respectively.

Quantitative analysis

In addition, a search for extended deletions and duplications of the *PAH* gene using the SALSA MLPA P055-050R PAH probemix set (MRC-Holland, the Netherlands) was carried out. The response was submitted according to the manufacturer's protocols. The data obtained were analyzed using the Coffalyser V8 software provided by the MLPA developer for accurate quantitative analysis.

Direct automatic sequencing sanger method

The nucleotide sequence of the *PAH* gene was determined by Sanger direct automatic sequencing of the PCR product from upstream and downstream primers. Fragments obtained through amplification by the PCR method were used as a sequencing template. Sequencing was carried out according to the manufacturer's protocol on the ABI Prism 3100 device (Applied Biosystems).

Results

Next generation sequencing

The results of an analysis of DNA samples from 259 patients who had one variant found in the *PAH* gene during previous studies, showed that 222 patients had a second variant with the first one confirmed. 37 patients had the first variant confirmed with no second variant revealed. 34 patients with no variants according to previous studies included 23 patients with *PAH* gene variants found: two variants in 15 patients, one variant in 8 patients. Two *PTS* gene variants were found in 5 patients. 6 patients had no revealed variants in the genes studied. 327 chromosomes without variants revealed were analyzed. According to the NGS method, *PAH* gene variants were found in 260 (10.3% of 2530 chromosomes) chromosomes, while *PTS* gene variants were found in 10 (0.4%) chromosomes.

110 rare variants were identified in the genes studied, including 104 – in the *PAH* gene, 6 – in the *PTS* gene. Ten variants in the *PAH* gene were reported for the first time. No variants were found in *GCH1*, *PCBD1*, *QDPR*, *SPR* and *DNAJC12* genes. All variants and their frequencies for the patient sample are given in Tables 1 and 2. For *PAH* gene variants, the calculation of allelic frequencies was performed on a sample of 2520 chromosomes, excluding chromosomes with *PTS* gene variants.

Quantitative analysis

The panel next generation sequencing method does not allow detecting major alterations and variants in untranslated regions. All the probands with no diagnosis confirmed by the NGS method were searched for gross deletions and duplications in the *PAH* gene using the MRC-Holland MLPA analysis method. During the study 45 probands had only one variant found, 6 probands had no variants found. Thus, quantitative analysis was required for 51 DNA samples of the probands, but the DNA of 13 of them was not suitable for MLPA analysis. Thus, the analysis was carried out for DNA samples of 38 probands. Among them, 10 patients with one *PAH* gene variant had gross deletions revealed, 28 patients had no gross deletions. Deletions in exons 1 (on 1 chromosome), 3 (on 4 chromosomes) and 5 (on 5 chromosomes) of the *PAH* gene were revealed. The deletions found and their frequencies are shown in Table 1. It should be noted that the most common variant of deletion of the exon 5 is included in 25 frequent *PAH* gene variants (EX5del4154ins268). Another variant of exon 5 deletion was revealed using the MLPA method. This method does not allow determining the exact boundaries of the deletions found. The gross deletions in the specified exons are described in the literature, but the matching of these variants with the deletions revealed in this study is yet to be confirmed (Birk Moller et al. 2007; Jeannesson-Thivisol et al. 2015; Trujillano et al. 2014).

Direct sanger sequencing

The next generation sequencing technology Ampliseq™ does not allow detecting deletions or duplications in homopolymer regions of a coding sequence containing more than 4 identical nucleotide residues. Exons 3, 6, 7 and 11 of the *PAH* gene contain homopolymer sequences of 5 nucleotides. For 35 patients with one variant and 6 patients with no variants, the analysis of the mentioned exons was carried out by the Sanger direct automatic sequencing method with

Table 1 Variants identified in the PAH gene

Position in cDNA	Position in protein	Number of chromosomes	Allele frequency ($N = 2530$),%	Pathogenicity of the variant[a]	PAH residual activity, %[b]
c.58C>T	p.Gln20*	2	0,08	P	–
c.60+5G>A	IVS1+5G>A	1	0,04	P	–
c.60+5G>T	IVS1+5G>T	8	0,32	P	–
c.169–13T>G	IVS2-13T>G	6	0,24	P	–
c.116_118del	p.Phe39del	2	0,08	P	–
c.158G>A	p.Arg53His	1	0,04	P	79
c.165T>G	p.Phe55Leu	1	0,04	P	–
c.165del	p.Phe55Leufs*6	10	0,40	P	–
c.194T>C	p.Ile65Thr	2	0,08	P	33
c.204A>T	p.Arg68Ser	2	0,08	P	25
c.205C>A	**p.Pro69Thr**	**1**	**0,04**	**LP**	–
c.248del	**p.Leu83Trpfs*6**	**5**	**0,20**	**P**	–
c.266dup	p.Ala90Cysfs*12	1	0,04	P	–
c.275C>T	p.Thr92Ile	2	0,08	P	76
c.311C>A	p.Ala104Asp	9	0,36	P	27
c.442–1G>A	IVS4-1G>A	3	0,12	P	–
c.355C>T	p.Pro119Ser	1	0,04	LP	–
c.395C>T	p.Ala132Val	1	0,04	LP	–
c.398_401del	p.Asn133Argfs*61	2	0,08	P	–
c.428A>T	p.Asp143Val	1	0,04	LP	–
c.434A>T	p.Asp145Val	2	0,08	P	–
c.436C>T	p.His146Tyr	1	0,04	LP	–
c.461A>T	p.Tyr154Phe	1	0,04	LP	–
c.472C>T	p.Arg158Trp	3	0,12	P	2
c.512G>A	p.Gly171Glu	1	0,04	LP	–
c.526C>T	p.Arg176*	3	0,12	P	–
c.529G>C	p.Val177Leu	1	0,04	P	–
c.533A>G	p.Glu178Gly	2	0,04	P	39
c.559T>C	p.Trp187Arg	6	0,26	P	–
c.563G>A	p.Gly188Asp	4	0,16	P	–
c.569T>C	p.Val190Ala	1	0,04	P	40
c.592_613del	p.Tyr198Serfs*136	2	0,08	P	–
c.611A>G	p.Tyr204Cys	3	0,12	P	–
c.620A>G	p.Asn207Ser	1	0,04	LP	–
c.631C>A	p.Pro211Thr	3	0,08	P	72
c.638T>C	p.Leu213Pro	1	0,04	P	–
c.648C>G	p.Tyr216*	3	0,12	P	–
c.665A>G	p.Asp222Gly	2	0,08	LP	–
c.665A>T	p.Asp222Val	1	0,04	LP	–
c.671T>C	p.Ile224Thr	1	0,04	P	–
c.674C>T	p.Pro225Leu	1	0,04	LP	–
c.673C>A	p.Pro225Thr	8	0,32	P	–
c.688G>A	p.Val230Ile	7	0,28	P	63
c.695A>C	**p.Gln232Pro**	**1**	**0,04**	**LP**	–
c.710G>T	**p.Cys237Phe**	**1**	**0,04**	**VUS**	–
c.712A>C	p.Thr238Pro	3	0,12	LP	–
c.721C>T	p.Arg241Cys	8	0,32	P	57
c.722G>A	p.Arg241His	5	0,20	P	23
c.730C>T	p.Pro244Ser	1	0,04	LP	–
c.734T>C	p.Val245Ala	2	0,08	P	50
c.755G>A	p.Arg252Gln	1	0,04	P	3
c.773T>G	p.Leu258Arg	1	0,04	LP	–
c.809G>A	p.Arg270Lys	1	0,04	P	11
c.826A>G	p.Met276Val	1	0,04	LP	–
c.837del	p.Glu280Asnfs*61	1	0,04	P	–
c.842+1G>A	IVS7+1G>A	7	0,28	P	–
c.842+5G>A	IVS7+5G>A	2	0,12	P	–
c.913-7A>G	IVS8-7A>G	2	0,08	P	–
c.913-1G>A	IVS8-1G>A	1	0,04	P	–
c.869A>T	p.His290Leu	1	0,04	LP	–
c.889C>T	p.Arg297Cys	1	0,04	LP	–
c.890G>A	p.Arg297His	1	0,04	P	39
c.874C>T	p.Pro292Ser	1	0,04	LP	–
c.896T>G	p.Phe299Cys	3	0,12	P	3

Table 1 (continued)

Position in cDNA	Position in protein	Number of chromosomes	Allele frequency ($N = 2530$),%	Pathogenicity of the variant[a]	PAH residual activity, %[b]
c.899C>A	**p.Ala300Asp**	**1**	**0,04**	**LP**	–
c.912+1G>C	IVS8+1G>C	1	0,04	P	–
c.970-1G>C	**IVS9-1G>C**	**1**	**0,04**	**P**	–
c.932T>C	p.Leu311Pro	7	0,28	P	10
c.934G>A	**p.Glu312Ser**	**1**	**0,04**	**LP**	–
c.960G>C	p.Lys320Asn	4	0,16	P	–
c.1066-14C>G	IVS10-14C>G	2	0,08	P	–
c.1066-3C>T	IVS10-3C>T	5	0,20	P	–
c.982A>C	**p.Thr328Pro**	**1**	**0,04**	**LP**	–
c.992T>C	p.Phe331Ser	2	0,08	P	–
c.1024del	p.Ala342Hisfs*58	1	0,04	P	–
c.1028A>G	p.Tyr343Cys	1	0,04	LP	–
c.1036G>A	p.Gly346Arg	1	0,04	LP	–
c.1042C>G	p.Leu348Val	4	0,16	P	25
c.1049C>A	p.Ser350Tyr	7	0,28	P	–
c.1055del	p.Gly352Valfs*48	1	0,04	P	–
c.1065+1G>A	IVS10+1G>A	4	0,12	P	–
c.1068C>A	p.Tyr356*	1	0,04	P	–
c.1089del	p.Lys363Asnfs	2	0,08	P	–
c.1114A>T	p.Thr372Ser	3	0,12	LP	–
c.1150C>T	p.Pro384Ser	1	0,04	P	76
c.1157A>G	p.Tyr386Cys	11	0,44	P	–
c.1159T>C	p.Tyr387His	3	0,12	P	–
c.1162G>A	p.Val388Met	2	0,08	P	28
c.1183G>C	p.Ala395Pro	1	0,04	P	16
c.1197A>T	p.Val399Val	3	0,12	P	–
c.1196T>C	p.Val399Ala	1	0,04	LP	–
c.1199+1G>C	IVS11+1G>C	5	0,20	P	–
c.1217T>C	p.Ile406Thr	1	0,04	LP	–
c.1223G>A	p.Arg408Gln	4	0,16	P	41
c.1238G>C	p.Arg413Pro	3	0,12	P	11
c.1243G>A	p.Asp415Asn	1	0,04	P	72
c.1246C>A	p.Pro416Thr	1	0,04	LP	–
c.1249T>A	p.Tyr417Asn	1	0,04	LP	–
c.1251C>A	**p.Tyr417***	**1**	**0,04**	**P**	–
c.1262T>C	p.Ile421Thr	1	0,04	P	–
c.1289T>C	p.Leu430Pro	2	0,08	P	–
c.1301C>A	p.Ala434Asp	1	0,04	P	9
c.1304A>T	**p.Asp435Val**	**2**	**0,08**	**VUS**	–
c.1357_1359+2del	p.*453Proext*33	1	0,04	LP	–
c.(?_-1)_(60+1_61–1)del	Ex1del	1	0,04	P	–
c.(168+1_169–1)_(352+1_353–1)del	Ex3del	4	0,16	P	–
c.(441+1_442–1)_(509+1_510–1)del	Ex5del	5	0,20	P	–

[a] *P* pathogenic variant, *LP* likely pathogenic variant, *VUS* variant of uncertain significance; [b] According to BioPKU.org database (BIOPKU 2006–2019); The variants not previously described are highlighted in bold

the purpose of detection of possible changes in homopolymer regions. 41 patients examined had no changes revealed in the nucleotide sequence of exons 3, 6, 7, and 11 of the *PAH* gene.

Of the 293 patients examined, 2 variants were found in 252 (82.3%), including 247 patients with the variants in the *PAH* gene and 5 (1.7%) patients with the variants in the *PTS* gene. One *PAH* gene variant was revealed in 35 (12.0%) probands. Variants in the genes studied were not found in 6 (2.1%) patients.

Discussion

Large-scale genotyping of patients with PKU and HPA using modern methods of DNA diagnostics was conducted for the first time in Russia. DNA samples of 293 unrelated probands, which were previously searched for 25 most frequent *PAH* gene variants, were studied according to the program of genotyping of patients with "phenylketonuria" and "hyperphenylaninemia" diagnoses. Based on the genotyping results, 259 of them had one variant found and 34 patients had

no variants found. The study found that 247 probands had two *PAH* gene variants and 35 probands had one *PAH* gene variant. Another 5 probands had two *PTS* gene variants found and 6 probands had no variants detected.

PAH gene variants

104 rare variants were identified in the *PAH* gene, including 72 pathogenic variants, 29 likely pathogenic variants and 3 variants of uncertain significance. The "phenylketonuria" diagnosis is considered to be confirmed if the patient has two pathogenic variants, or one pathogenic and one likely pathogenic variants, or two likely pathogenic variants, in the *PAH* gene. If a patient with one pathogenic or likely pathogenic variant has a variant of uncertain significance or two uncertain significance variants – the "phenylketonuria" diagnosis is considered to be not confirmed. As a result, the "phenylketonuria" diagnosis was confirmed for 245 patients (83.6% of 293 probands). This diagnosis has not been confirmed yet for the 2 patients who had two variants in the *PAH* gene, as they had a variant of uncertain significance revealed. Variants were not found in 6 patients, while 35 patients had only 1 variant found in the *PAH* gene, which does not allow either to confirm or disprove the diagnosis.

Ten variants were not previously described in the literature: p.Pro69Thr, p.Leu83Trpfs*9, p.Gln232Pro, p.Cys237Phe, p.Ala300Asp, p.Gly312Ser, p.Thr328Pro, p.Tyr417*, p.Asp435Val and IVS9-1G>C. Variants p.Leu83Trpfs*9, p.Tyr417* and IVS9-1G>C are pathogenic, two variants p.Cys237Phe and p.Asp435Val are classified as variants of uncertain significance. The remaining variants are classified as likely pathogenic. Pathogenicity criteria for this variants are presented in Table 3.

The BioPKU.org database contains data on the residual activity of the PAH enzyme for some variants of the *PAH* gene. According to the degree of residual activity of the enzyme, all variants can be divided into "severe" (residual activity below 10%) and "mild" (residual activity above 10%). In this study, 104 identified variants included 21 mild and 17 severe variants. The BioPKU.org database contains no data on the residual activity of the PAH enzyme for 66 variants. In Russia, mutant alleles of the *PAH* gene, revealed as a result of search of 25 most frequent *PAH* gene variants, contain both severe variants (74% of all PKU alleles), and mild variants (12%) (Gundorova et al. 2019). According to the results of the work performed, the set of patients with PKU, without taking into account the probands having variants in the *PTS* gene (2520 chromosomes), had 62 alleles with mild variants and 56 alleles with severe variants revealed, while the residual activity of the enzyme is unknown for 179 alleles. As a result, the share of severe variants in a set of 1260 probands amounted to 75.9%, and the share of mild ones – 14.4%. The share of patients potentially responsive to the BH4 therapy increased by 2.5%, and the share of probands potentially unresponsive to therapy – by 2.3% (Fig. 1).

In the analysis of 25 most frequent variants, at least one "mild" variant was found in 22.6% of cases. The analysis using the NGS method revealed "mild" variants in 2.4% of cases in addition, however, in 9.7% of cases the residual activity of the enzyme is unknown. This does not allow to predict the potential BH4 sensitivity by the genotype. Loading testing is required to determine the sensitivity to BH4 therapy for patients with these variants. If the purpose of the DNA analysis is only to determine the BH4 sensitivity, it is advisable to search for 25 frequent variants. In case if the sensitivity could not be determined, it is suggested to carry out BH4 loading test. The use of next generation sequencing for sensitivity prediction does not justify time and economic costs, but it is suitable for determining the genotype of a patient in genetic research and for family consulting.

BH4-deficient forms

6 variants in the *PTS* gene were found during this study including 2 pathogenic variants, one likely pathogenic variant and 3 variants of uncertain significance. These variants lead to the development of BH4-deficient hyperphenylalaninemia A. The most commonly encountered variant of uncertain significance was p.Arg9His, found in the population of East Asia (Chiu et al. 2012). This variant was found in one proband in homozygous state. Disorder in the synthesis of pyruvoyltetrahydrobiopterin synthetase was confirmed by

Table 2 Variants identified in the *PTS* gene

Exon	Position in cDNA	Position in protein	Number of chromosomes	Allele frequency ($N = 2530$),%	Pathogenicity of the variant[a]
1	c.26G>A	p.Arg9His	3	0,12	VUS
2	c.94A>G	p.Ser32Gly	1	0,04	VUS
2	c.108C>G	p.Asn36Lys	1	0,04	VUS
2	c.146A>G	p.His49Arg	1	0,04	LP
4	c.216T>A	p.Asn72Lys	2	0,08	P
6	c.317C>T	p.Thr106Met	2	0,08	P

[a] *P* pathogenic variant, *LP* likely pathogenic variant, *VUS* variant of uncertain significance

Table 3 Pathogenicity of the new variant in the *PAH* gene

Variant	Criteria for Pathogenicity[a]	Pathogenicity of the variant[b]
p.Pro69Thr	PS3, PM2, PP2, PP3	LP
p.Leu83Trpfs*6	PVS1, PM2, PP3	P
p.Gln232Pro	PM2, PM5, PP2, PP3, PP4	LP
p.Cys237Phe	PM2, PP2, PP3	VUS
p.Ala300Asp	PM2, PM5, PP2, PP3, PP4	LP
IVS9-1G>C	PVS1, PM2, PP3	P
p.Glu312Ser	PM2, PM5, PP2, PP3, PP4	LP
p.Thr328Pro	PS3, PM2, PP2, PP3, PP4	LP
p.Tyr417*	PS3, PM2, PM4, PP3, PP4	P
p.Asp435Val	PM2, PP2, PP3	VUS

[a] The clinical significance (pathogenesis) of the variants revealed was evaluated on the basis of recommendations for the interpretation of data received by next generation sequencing methods (Richards et al. 2015); [b] *P* pathogenic variant, *LP* likely pathogenic variant, *VUS* variant of uncertain significance

positive analysis of pterins in urine conducted for the proband; therefore, this variant may be a possible cause of the pathology. Variants p.Ser32Gly, p.Asn72Lys and p.Thr106Met had been previously discovered in the course of routine diagnostics from 2005 to 2006 in the DNA-diagnostics laboratory of Federal State Funded Research Institution "Research Centre of Medical Genetics" (Stepanova et al. 2006). However, variant P.Ser32Gly has been described earlier in one article only and is classified as a variant of uncertain significance. In this study, as well as in Stepanova A.A.'s paper, this variant was found in patients with a pathogenic variant (Stepanova et al. 2006). Variant p.Ser32Gly may be a possible cause of a pathology. Variant p.Asn36Lys was described by Zekanowski, and is a variant of uncertain significance (Zekanowski et al. 1998). The likely pathogenic variant p.His49Arg had not been found earlier in any group of HPA patients in Russia; in this study, it was discovered in a patient with a pathogenic variant. According to the results of the study, the diagnosis "HPABH4A" was confirmed for two patients. The diagnosis was not confirmed for three patients, since their genotype included a variant of uncertain significance.

BH4-deficient hyperphenylalaninemia develops due to variants in genes *PTS*, *GCH1*, *PCBD1*, *QDPR*, *SPR*. These forms make up 1–2% of all the HPA cases globally. The most frequent of them – HPABH4A and HPABH4C – are caused by variants in genes *PTS* and *QDPR*, respectively. The share of *PTS*-caused forms accounts for 65.3% of all BH4-deficient HPA cases, *QDPR* – 24.9% of all HPABH4 cases (Blau 2016).

The incidence of HPA diseases in the world, in particular in Russia, was compared in this study. In the course of a literature review, it was found that similar studies were conducted in China (Li et al. 2015; Wang et al. 2018). The NGS method

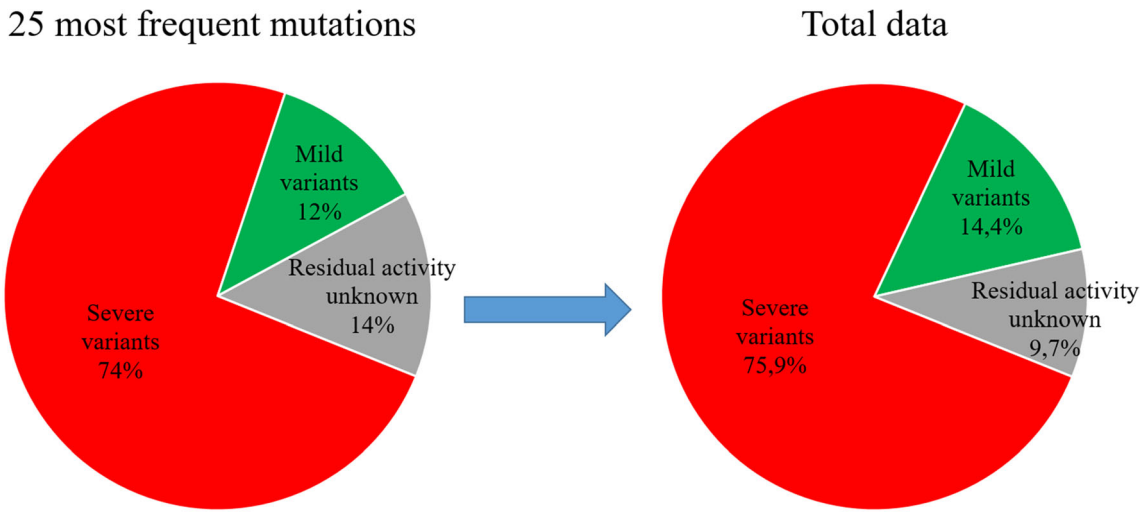

Fig. 1 Shares of severe and mild variants of the *PAH* gene (*N* = 1260 probands). The left chart shows the shares of severe and mild variants detected in the search for the 25 most frequent variants. The right chart shows the summary data obtained after the search of 25 most frequent variants, next generation sequencing and quantitative analysis of the *PAH* gene. Red color shows the shares of severe variants, green color - mild variants, gray region includes variants for which the residual enzyme activity is unknown

was used as the main detection method in these studies, as well as an additional method for detecting variants. According to Chinese scientists, the share of *PAH*-mediated HPA forms is 74.8% of all cases, BH4-deficient forms of HPA: HPABH4A (19.7%), HPABH4C (0.7%) and HPABH4B with 0.1% incidence, other rare HPA forms were not found. In Russia, BH4-deficient HPA is represented mainly by HPABH4A (Stepanova et al. 2006). According to the results of this study, its share amounts to 0.4% of all HPA cases. According to the findings of this study, the incidence of rare HPA forms in Russia is lower than in Europe (1–2%) (Blau 2016).

In 2017, the connection between *DNAJC12* gene variants and HPA was established. This state was called "non-BH4-deficient HPA". Consequently, the *DNAJC12* gene was included in the custom NGS panel. This gene has not yet been studied in Russia. No variants were found in this gene. In this study, no variants were also revealed in *GCH1*, *PCBD1*, *QDPR* and *SPR* genes.

A second variant in the *PAH* gene was not found in 35 probands. This may be due to the presence of variants in the non-coding and regulatory regions of the *PAH* gene. Also, no variant was detected in 6 probands. This may be due to the presence of two variants in the non-coding and regulatory regions of *PAH*, *PTS*, *GCH1*, *PCBD1*, *QDPR*, *SPR* and *DNAJC12* genes. It can be assumed that there are other genes, whose relation to HPA development has not yet been established.

All patients who participated in the genotyping program underwent biochemical screening, re-screening, and Phe level change analysis by physicians. All these examinations showed an increase in the level of Phe in the blood. This eliminates the likelihood of a false HPA diagnosis.

Conclusions

According to the study results, a spectrum of variants leading to hyperphenylalaninemia in the territory of the Russian Federation was established. The analysis was conducted using the next generation sequencing technology. This method has allowed analyzing a large set of patients for the presence of variants in genes *PAH*, *PTS*, *GCH1*, *PCBD1*, *QDPR*, *SPR* and *DNAJC12* within a short period of time. It allowed confirmation of diagnosis in an additional 247 patients with *PAH* deficiency and 5 patients with *PTS* deficiency and in an additional 35 probands one *PAH* gene variant was found. The data obtained helped reveal rare variants of the *PAH* gene and establish the percentage of tetrahydrobiopterin-dependent HPA forms in Russia. The share of rare forms amounted to 0.4% of all HPA forms.

This study will help optimize the existing diagnostics protocol. As a first stage, a search for the 25 most frequent variants in the *PAH* gene is advised. If one frequent variant in the *PAH* gene is detected, a search for the second variant by the *PAH* gene Sanger sequencing is suggested as a second stage. If frequent variants are not revealed, genotyping by the NGS panel is advised. If the diagnosis is not confirmed, a search for gross insertions and deletions by the MLPA method is recommended as a third stage.

Acknowledgements The authors are grateful for the participation of all patients in this study. We would like to thank clinical genetic specialists from all over the country for their cooperation.

Funding The research was carried out within the state assignment of Ministry of Science and Higher Education of the Russian Federation.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

References

Anikster Y, Haack TB, Vilboux T, Pode-Shakked B, Thöny B, Shen N, Guarani V, Meissner T, Mayatepek E, Trefz FK, Marek-Yagel D, Martinez A, Huttlin EL, Paulo JA, Berutti R, Benoist JF, Imbard A, Dorboz I, Heimer G, Landau Y, Ziv-Strasser L, Malicdan MCV, Gemperle-Britschgi C, Cremer K, Engels H, Meili D, Keller I, Bruggmann R, Strom TM, Meitinger T, Mullikin JC, Schwartz G, Ben-Zeev B, Gahl WA, Harper JW, Blau N, Hoffmann GF, Prokisch H, Opladen T, Schiff M (2017) Biallelic mutations in DNAJC12 cause Hyperphenylalaninemia, dystonia, and intellectual disability. Am J Hum Genet 100:257–266. https://doi.org/10.1016/j.ajhg.2017.01.002

BIOPKU (2006-2019) http://www.biopku.org

Birk Moller L et al (2007) Low proportion of whole exon deletions causing phenylketonuria in Denmark and Germany. Hum Mutat 28:207. https://doi.org/10.1002/humu.9481

Blau N (2016) Genetics of phenylketonuria: then and now. Hum Mutat 37:508–515. https://doi.org/10.1002/humu.22980

Bonafe L, Thony B, Penzien JM, Czarnecki B, Blau N (2001) Mutations in the sepiapterin reductase gene cause a novel tetrahydrobiopterin-dependent monoamine-neurotransmitter deficiency without hyperphenylalaninemia. Am J Hum Genet 69:269–277. https://doi.org/10.1086/321970

Chiu YH, Chang YC, Chang YH, Niu DM, Yang YL, Ye J, Jiang J, Okano Y, Lee DH, Pangkanon S, Kuptanon C, Hock NL, Chiong MA, Cavan BV, Hsiao KJ, Liu TT (2012) Mutation spectrum of and founder effects affecting the PTS gene in east Asian populations. J Hum Genet 57:145–152. https://doi.org/10.1038/jhg.2011.146

Dhondt JL, Farriaux JP, Boudha A, Largilliere C, Ringel J, Roger MM, Leeming RJ (1985) Neonatal hyperphenylalaninemia presumably caused by guanosine triphosphate-cyclohydrolase deficiency. J Pediatr 106:954–956

Dudesek A, Roschinger W, Muntau AC, Seidel J, Leupold D, Thony B, Blau N (2001) Molecular analysis and long-term follow-up of

patients with different forms of 6-pyruvoyl-tetrahydropterin synthase deficiency. Eur J Pediatr 160:267–276

Grobe H, Bartholome K, Milstien S, Kaufman S (1978) Hyperphenylalaninaemia due to dihydropteridine reductase deficiency. Eur J Pediatr 129:93–98

Gundorova P, Stepanova AA, Kuznetsova IA, Kutsev SI, Polyakov AV (2019) Genotypes of 2579 patients with phenylketonuria reveal a high rate of BH4 non-responders in Russia. PLoS One 14:e0211048. https://doi.org/10.1371/journal.pone.0211048

Jeannesson-Thivisol E, Feillet F, Chéry C, Perrin P, Battaglia-Hsu SF, Herbeth B, Cano A, Barth M, Fouilhoux A, Mention K, Labarthe F, Arnoux JB, Maillot F, Lenaerts C, Dumesnil C, Wagner K, Terral D, Broué P, de Parscau L, Gay C, Kuster A, Bédu A, Besson G, Lamireau D, Odent S, Masurel A, Guéant JL, Namour F (2015) Genotype-phenotype associations in French patients with phenylketonuria and importance of genotype for full assessment of tetrahydrobiopterin responsiveness. Orphanet J Rare Dis 10:158. https://doi.org/10.1186/s13023-015-0375-x

Li N, Jia H, Liu Z, Tao J, Chen S, Li X, Deng Y, Jin X, Song J, Zhang L, Liang Y, Wang W, Zhu J (2015) Molecular characterisation of phenylketonuria in a Chinese mainland population using next-generation sequencing. Sci Rep 5:15769. https://doi.org/10.1038/srep15769

Richards S et al (2015) Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 17:405–424. https://doi.org/10.1038/gim.2015.30

Stepanova AA, Polyakov AV, Tverskaya SM (2006) Molekulyarno-geneticheskoye issledovaniye genov tetrogidrobiopterinovogo obmena QDPR i PTS v gruppe rossiyskikh bol'nykh fenilketonuriyey [in russ]. Med Genet 9:3–8

Thony B et al (1998) Hyperphenylalaninemia with high levels of 7-biopterin is associated with mutations in the PCBD gene encoding the bifunctional protein pterin-4a-carbinolamine dehydratase and transcriptional coactivator (DCoH). Am J Hum Genet 62:1302–1311. https://doi.org/10.1086/301887

Trujillano D, Perez B, González J, Tornador C, Navarrete R, Escaramis G, Ossowski S, Armengol L, Cornejo V, Desviat LR, Ugarte M, Estivill X (2014) Accurate molecular diagnosis of phenylketonuria and tetrahydrobiopterin-deficient hyperphenylalaninemias using high-throughput targeted sequencing. Eur J Hum Genet 22:528–534. https://doi.org/10.1038/ejhg.2013.175

van Wegberg AMJ, MacDonald A, Ahring K, Bélanger-Quintana A, Blau N, Bosch AM, Burlina A, Campistol J, Feillet F, Giżewska M, Huijbregts SC, Kearney S, Leuzzi V, Maillot F, Muntau AC, van Rijn M, Trefz F, Walter JH, van Spronsen FJ (2017) The complete European guidelines on phenylketonuria: diagnosis and treatment. Orphanet J Rare Dis 12:162. https://doi.org/10.1186/s13023-017-0685-2

Wang R, Shen N, Ye J, Han L, Qiu W, Zhang H, Liang L, Sun Y, Fan Y, Wang L, Wang Y, Gong Z, Liu H, Wang J, Yan H, Blau N, Gu X, Yu Y (2018) Mutation spectrum of hyperphenylalaninemia candidate genes and the genotype-phenotype correlation in the Chinese population. Clin Chim Acta 481:132–138. https://doi.org/10.1016/j.cca.2018.02.035

Zekanowski C, Nowacka M, Sendecka E, Sowik M, Cabalska B, Bal J (1998) Identification of mutations causing 6-Pyruvoyl-tetrahydrobiopterin synthase deficiency in polish patients with variant Hyperphenylalaninemia. Mol Diagn 3:237–239 https://doi.org/10.1016/S1084-8592(98)80047-7

Zurfluh MR et al (2008) Molecular genetics of tetrahydrobiopterin-responsive phenylalanine hydroxylase deficiency. Hum Mutat 29:167–175. https://doi.org/10.1002/humu.20637