MARIA SCURTUL, CHIRIL BOICIUC, DANIELA BLĂNIŢĂ, TAMARA CROITORI,
NICOLETA LAZARI, NATALIA UŞURELU*

MOLECULAR GENETICS IN PHENYLKETONURIA IN REPUBLIC OF MOLDOVA (2018- 2019)

Institute of Mother and Child, Chisinau, Republic of Moldova

REZUMAT

SPECIFICUL MOLECULAR GENETIC ÎN FENILCETONURIE ÎN REPUBLICA MOLDOVA (anii 2018-2019)

Introducere: Fenilcetonuria (*PKU*) este o patologie autosomal recesivă, cauzată de deficienţa enzimei fenilalanin hidroxilaza, care duce la acumularea fenilalaninei (*Phe*) serice, cauzând retard mental.

Scopul lucrării: Diagnosticarea *PKU* în baza screeningului neonatal, organizat în perioada 2018-2019, precum şi cercetarea, la nivel molecular genetic, a pacienţilor identificaţi pentru a stabili mutaţiile cauzatoare de boală.

Material şi metode: În perioada 2018-2019 au fost supuşi screeningului 62943 de nou-născuţi din Moldova (rata de screening 96,16%), în baza metodei fluorometrice. La 17 copii a fost detectată *PKU*, iar 11 au fost testaţi genetic prin metoda *PCR/RFLP* la 6 mutaţii în gena *PAH* (*p.R408W, p.P281L, p.L48S, .p.R158Q, p.R252W, p.R261Q*) şi secvenţierea exonilor 2,3,5,6-8,10-13 ai genei *PAH* după metoda Sanger.

Rezultate: Analiza molecular-genetică prin *PCR/RFLP* a permis detectarea mutaţiilor în 45,45% de cazuri. Cel mai frecvent genotip a fost *p.R408W/p.L48S*, stabilit la 3 pacienţi. Analiza suplimentară prin metoda Sanger a permis detectarea genotipului complet în restul cazurilor, determinând mutaţii de tip splicing (*IVS7+4A>G, IVS11+7T>C, IVS12+1G>A*), missense (*p.A342T*) şi nonsense (*p.R111**).

Concluzii: În urma studiului a fost detectat un spectru de mutaţii specifice cohortei PKU din Moldova, datorită metodelor eficiente de screening şi diagnostic.

Cuvinte-cheie: fenilcetonurie, gena fenilalanin hidroxilazei, screening neonatal.

РЕЗЮМЕ

МОЛЕКУЛЯРНО-ГЕНЕТИЧЕСКАЯ СПЕЦИФИКА ФЕНИЛКЕТОНУРИИ В РЕСПУБЛИКЕ МОЛДОВА (2018- 2019)

Введение: Фенилкетонурия (ФКУ) – это генетическое заболевание с аутосомно рецессивной моделью наследования обусловленное дефицитом фенилаланин- гидроксилазы. Дисфункция данного энзима приводит к накоплению фенилаланина в крови и последующим дегенеративным процессам в нервной системе.

Цель: Целью исследования была диагностика пациентов с ФКУ с помощью неонатального скрининга, в период 2018-2019 годов, а так же проведение молекулярно- генетического изучения пациентов для идентификации мутаций в гене *PAH*.

Материалы и методы: В период 2018-2019 годов, неонатальному скринингу подверглись 62943 новорождённых (доля 96,16%), по средством флуорометрического метода. У 17 детей была выявлена ФКУ, 11 из которых были подвержены ПЦР/ПДРФ анализу с целью выявления 6 частых мутаций в гене *PAH* (*p.R408W, p.P281L, p.L48S, p.R158Q, p.R252W, p.R261Q*) и секвенированию экзонов 2,3,5,6-8,10-13 гена *PAH* при помощи капилярного электрофореза по методике Sanger.

Результаты: В результате проведения скрининга, были выявлены 376 подозреваемых случаев ФКУ, у которых показатели фенилаланина были больше 3 мг/дЛ, Образцы ДНК были получены у 11 из 17 подтвержденных пациентов. Молекулярно- генетическое исследование при помощи ПЦР/ПДРФ, позволило выявить мутации в 45,45% случаев. Самым распространённым генотипом оказался *p.R408W/ p.L48S*, выявленный у 3 пациентов. Пополнение исследования результатами секвенирования кодирующей части гена *PAH* позволило определить полный генотип в остальных случаях.

Ключевые слова: фенилкетонурия, ген фенилаланин-гидроксилазы, фенилаланин.

Introduction. Phenylketonuria (*PKU*) is an autosomal recessive disorder, caused by phenylalanine hydroxylase deficiency. Phenylalanine hydroxylase enzyme is encoded by *PAH* gene located on the long arm of the chromosome 12. This enzyme has tetrahydrobiopterin (*BH4*) as endogenous cofactor, and together they lead to conversion of phenylalanine (*Phe*) into tyrosine. Due to presence of mutations in *PAH* gene, the enzyme is not able to work properly, so that the level of *Phe* increases. The main symptom of *PKU* is mental retardation, associated with eczema, lower pigmentation, and others [1].

In many studies there was observed a direct correlation between enzyme dysfunction and the level of *Phe* blood accumulation, determining the nervous system damage[2]. There are known "*major*" and "*minor*" mutations in *PAH* gene associated with different forms of severity[3]. The typical and the most severe is classical *PKU*, when phenylanine hydroxylase protein is non-working, contributing to high *Phe* concentration in blood (>1200 μmol/L). Mild *PKU*, is a form that is characterized by an increased activity of phenylanine hydroxylase in comparison with classic form and moderate severity of disease manifestation, having lower *Phe* blood levels. Those forms particularly require prompt treatment in order to prevent severe mental retardation. The third type- Hyperphenylalaninemia can be described by slightly elevated *Phe* blood levels (<600 μmol/L), and lower risks associated with mental retardation. There were reported missense, nonsense, splicing mutations, and deletions in *PAH* gene, causing *PKU*. In addition, there are atypical forms of *PKU*, independent of *PAH* gene, encountered in approximately 2% of cases, caused by *BH4* dependent enzymes (PTPS, SR, DHPR, GTPCH)[4].

It is extremely important to detect PKU before the onset of the disease to start the treatment and to prevent mental retardation and other outcomes induced by the disease [2].

In the Republic of Moldova, the neonatal screening was performed for the first time in 1989, started as a pilot program that was extended to national level, now covering 95-98% of newborns[5]. Previous assessment showed that the frequency of PKU in Republic of Moldova is 1 to 7325 newborns [6].

The aim of the study was to identify the pathologic genotypes that caused abnormalities in *Phe* metabolism in *PKU* patients identified during 2018-2019.

Material and Methods. The DNA samples, collected from 11 newborns out of 17 with *Phe* blood level higher than 3 mg/dL, identified in 2018-2019 through neonatal screening were used as materials for the given study. The fluorometric method with McCaman&Robins modification was used as the main technique of *Phe* quantification, based on detection kits from Anilabsystems (Finland) for neonatal screening[6].

The first step in molecular genetic testing was checking the patient for 6 common mutations (p.R408W, p.P281L, p.L48S, p.R158Q, p.R252W, p.R261Q) by *PCR/RFLP*[7]. In those cases when full pathological genotype could not be detected by *PCR/RFLP*, Sanger Sequencing was used for 2,3,5-7,10-12 exons and adjacent at least 10-15 bp of intron region[8]. Sanger Sequencing was performed on ABI Genetic Analyzer equipment (USA), using BigDye Terminator v3.1 as sequencing chemistry. Capillary electrophoresis was performed using 50 cm array and POP7 as polymer[9].

The obtained raw data were analyzed using "SeqA 6" and "SeqScape 3"[10]. The resulted sequences of *PAH* gene were aligned to NG_008690.2 refference. Meanwhile, the identified genetic variants were described accordingly to HGVS nomenclature and using as refference NM_000277.3 sequence[11].

The next step was the estimation of variant pathogenicity, using *in silico* prediction algorithms suitable for missense variants (PolyPhen-2 and SIFT), for splicing mutations (NetGene2 and FSPLICE), and BIOPKU database (BIOPKU, http://www.biopku.org) for descriptions of known mutations in *PAH* gene. The PolyPhen-2 tool is based on the principle of sequence analysis, protein function and its structure, along with the phylogenetic conservation that allows prediction of the effect of the mutation on the amino acid sequence. At the same time, SIFT uses evolutionary conservation as the main policy for prediction of the effect on the protein structure[12]. The NetGene2 service is based on neural networks that detect the locations of splice sites [13]. The FSPLICE tools is a species- specific splice site predictor based on weight matrices model[12].

Results. The first part of the molecular aspect of the study was based on *PCR/RFLP* method that had a solid role in diagnosis of Moldovan *PKU* cohort. It is specific for detection of six mutations (*p.R408W, p.P281L, p.L48S, p.R158Q, p.R252W, p.R261Q*) that were frequently encountered in Moldovan population. In 11 patients that were taken into account in the research, *PCR/RFLP* led to revealing of six genotypes out of total (54,55%).

The *PCR/RFLP* analysis revealed that *p.R408W* was the most frequent mutation in the studied patients (45,45%), followed by *p.L48S* (13,64%). These two mutations made up the most frequent genotype in our *PKU* cohort *p.R408W/p.L48S*, detected in three patients (27,27%) during these two years. The p.R408W mutation in homozygous state was being observed in two patients from total of 11 (18,18%). The other six variants occurred in one patient at a time, without any repetition.

The study was replenished with data from Sanger Sequencing that contributed to the elevation of the detection rate to approximately 100%, by completing the existent genotypes with 11 mutations (*IVS7+4A>G, IVS11+7T>C, IVS12+1G>A, p.R111X, p.A342T, c.1065+97G>A, c.1065+245A>T, c.1066-236C>T, L385L, Q232Q, V245V*). Considering that Sanger Sequencing allows analysis of a region of DNA around 1000 bp, the results using this method brought diversity into the range of mutations. After the analysis using prediction tools (SIFT, PolyPhen) and BIOPKU database, five of them were classified as pathological. Other three were identified as splicing mutations (*IVS7+4A>G, IVS11+7T>C, IVS12+1G>A*) analyzed by NetGene2 and FSPLICE, affecting the splice site and making the protein shorter.

The five mutations analyzed with VEP showed little difference in results. According to SIFT, all variants received a score of 0, being classified as deleterious on the protein. At the same time, PolyPhen-2 attributed 4 out of 5 mutations a score of 1 (likely pathogenic), except *p.R261Q*, this variant receiving a score of 0.997 (probably pathogenic). Based on ClinVar archives, all the mutations are pathogenic or likely pathogenic, only *p.A342T* was not classified in any category because of lack of data (tab 2).

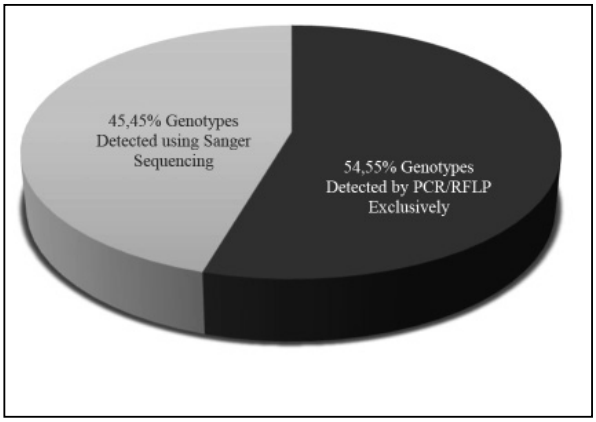

Fig. 1. Rate of molecular genetic tests in diagnostics of PKU by various methods

It is important to mention that not all the exons in the *PAH* gene have been sequenced due to lack of specific primers for exons 1, 4, 8, 9, and 13. But at the same time, in the studied cases, there were obtained the mutations inducing *PKU*, so having the suggestion that the pathological mutations are not located in the unanalyzed exons. After the analysis of pathogenicity, using various prediction tools and BIOPKU database, five of them were classified as pathologic. From those five pathological variants, one was identified as a nonsense mutation (*p.R111**), because of the severe structural modifications in the protein it produces, due to a premature stop codon in exon 3 (fig. 2). According to BIOPKU database, this variant was first reported in 1992, being a null mutation localized in regulatory domain.

A missense variant detected in a single patient from the study was *p.A342T*, localized in the exon 10 (fig. 2). Prediction tools as SIFT and PolyPhen-2 estimated its pathogenicity accordingly to the negative score (SIFT-0,0 (deleterious), PolyPhen-2- 1,0 (deleterious)) (tab.2). By consulting with BIOPKU database, it was demonstrated that this variant was already identified in *PKU* patients, being mentioned for the first time in 1992. This mutation affects the catalytic domain, but at the same time, the enzyme activity of *PAH* was demonstrated to be around 26%. Therefore, it can be concluded that this variant leads to a milder phenotype of *PKU*.

Due to Sequencing analysis of *PAH* gene, there were identified three different splicing mutations, that were the most frequent type of mutation in the given cohort. According to the fact that splicing variants are usually localized near exon-intron junction, it was not possible to use SIFT and PolyPhen-2 as prediction tools, only suitable for missense variants. To identify the effect on splicing of the given variants NetGene2 and FSPLICE tools were used, especially designed for detecting donor/acceptor regions that makes possible to check whether the splice site is lost or not. The analysis showed that all three mutations *IVS7+4A>G, IVS11+7T>C, IVS12+1G>A* are produced in donor/acceptor regions, resulting in deletion of splice site and consequently to a shorter protein with altered amino acid sequence, so it was possible to assume that those variants are pathological.

The *IVS12+1G>A* was identified in two patients in heterozygous state. This variant is a substitution of guanine with adenine, localized in intron 12 (fig.2), being reported in BIOPKU database as a pathologic mutation

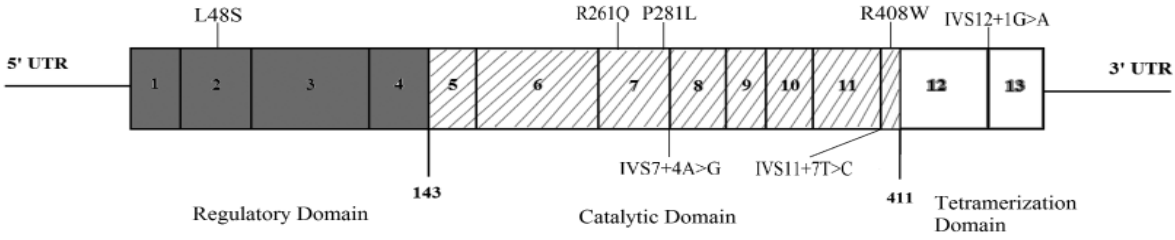

Figure 2. PAH gene structure and mutation localization

(tab. 3). Another splicing mutation identified in the studied *PKU* cohort was *IVS7+4A>G*, found in intron 7 (fig.2). The analysis of the sequence in NetGene2 and FSPLICE showed that this nucleotide substitution provokes splicing site disappearing, which suggested that this variant is a pathologic one. Moreover, *IVS7+4A>G* was reported in BIOPKU database as a null mutation, causing a severe form of *PKU*. Another identified splicing variant is *IVS11+7T>C*. This mutation brought some uncertainties in the study because the analysis with NetGene2 and FSPLICE showed no change in the splice site, caused by this mutation, and there were no records with it in BIOPKU database. However, it can be assumed that this variant is pathologic, due to the fact,

Table 1.

PAH gene mutations detected in PKU cohort 2018-2019

Patient number	Allele 1	Allele 2	Method of Detection
1	R408W	L48S	PCR/RFLP
2	R408W	P281L	PCR/RFLP
3	R408W	L48S	PCR/RFLP
4	IVS7+4A>G	IVS11+7T>C	Sequencing
5	R261Q	A342T	Combined
6	P281L	IVS12+1G>A	Combined
7	R408W	R408W	PCR/RFLP
8	R408W	R408W	PCR/RFLP
9	R408W	L48S	PCR/RFLP
10	R408W	R111*	Combined
11	R408W	IVS12+1G>A	Combined

Table 2.

Effect Prediction of Missense PAH Gene Variants Using Prediction Tools

HGVS Nomenclature	dbSNP Nomenclature	SIFT Score (interpretation)	Polyphen Score (interpretation)	ClinVar clinical significance
R408W	rs5030858	0 (deleterious)	1 (likely pathogenic)	Pathogenic
L48S	rs5030841	0 (deleterious)	1 (likely pathogenic)	Pathogenic
P281L	rs5030851	0 (deleterious)	1 (likely pathogenic)	Pathogenic/ likely pathogenic
R261Q	rs5030849	0 (deleterious)	0,997 (probably damaging)	Pathogenic
A342T	rs62507282	0 (deleterious)	1 (likely pathogenic)	Not provided

for the first, the patient shows high *Phe* blood levels, and additionally in BIOPKU database there were other splice mutations in intron 11, that are located further in the intron sequence (*ex.IVS11+17G>A, IVS11+20G>C*), showing that the splice site donor or acceptor is larger than it was thought to be. Despite this fact, a functional analysis should be performed in order to understand the full influence of the variant on the *PAH* enzyme structure.

2. The results showed that *p.R408W* and *p.L48S* mutations prevailed in PKU patients diagnosed in 2018-2019 in Moldova, along with a majored proportion of splicing mutations.

3. The prediction tools proved to be extremely important in gene mutations interpretation in order understand the form of PKU and to decide on the final necessity of the life- treating diet and prenatal diagnosis as well.

Table 3.

Splice Site Lost Prediction in Intron Variants of PAH Gene

Mutation	**NetGene2**		**FSPLICE**	
	Normal	Mutant	Normal	Mutant
IVS7+4A>G	GTGaGTACTG 0.93 confidence donor site	GTGgGTACTG 0.60 confidence donor site	GAACCGTGaGTA >90% donor site	GAACCGTGgGTA 0% donor site
IVS11+7T>C	GTGAGGtGGT 0.96 confidence donor site	GTGAGGcGGT 1.0 confidence donor site	GTAAGGTGAGGt >90% donor site	GTAAGGTGAGGc >90% donor site
IVS12+1G>A	gTAAGTAATT 0.95 confidence donor site	aTAAGTAATT 0.00 confidence donor site	TAACAgTAAGTA >90% donor site	TAACAaTAAGTA 0% donor site

Finally, all the included patients with *PKU* have been identified with two pathological mutations that caused the disease. However, using Sanger Sequencing, it was possible to detect six non- pathological variants additionally in our *PKU* patients. Three of them were silent mutations (*p.L385L, p.Q232Q, p.V245V*) localized in the coding region, without affecting the amino acid sequence, accordingly to "SecScape 3" program. The other three were identified in the deep intron region that cannot usually affect the normal splicing of the mRNA (*c.1065+97G>A, c.1065+245A>T, c.1066-236C>T*)[8].

Discussions. Given all that, during the experience of the lab, an optimal algorithm of PKU diagnostics was elaborated. This algorithm consists of a combination of this two techniques (PCR/RFLP, Sanger Sequencing) that demonstrated a good performance, cost- effectiveness while the detection rate was high. The complete genotyping of the PKU patients is essential for establishing the form of PKU and brings consent in diet necessity. Additionally, it is used for final acceptance of the disease by patient and prenatal tests in the future if asked for.

Conclusions

1. During the molecular genetic diagnostics, all the pathologic mutations have been detected in the studied group. The combination of *PCR/RFLP* and Sanger Sequencing was established as very successful approach, achieving high performance results.

Bibliography

1. Van Wegberg AMJ, MacDonald A, Ahring K, et al. The complete European guidelines on phenylketonuria: Diagnosis and treatment. *Orphanet J Rare Dis*. 2017;12(1):1-56. doi:10.1186/s13023-017-0685-2
2. Vliet D Van, Wegberg AMJ Van, Ahring K, et al. Can untreated PKU patients escape from intellectual disability ? A systematic review. 2018:1-6.
3. Scurtul Maria, Chiril Boiciuc, Daniela Blăniţă, Victoria Sacară NU. PHENOTYPE PREDICTION IN PHENYLKETONURIA PATIENTS FROM MOLDOVA BASED ON GENOTYPE DATA. *Bul Perinatol*. 2020.
4. Sumaily KM, Mujamammi AH. Phenylketonuria: A new look at an old topic, advances in laboratory diagnosis, and therapeutic strategies. *Int J Health Sci (Qassim)*. 2017;11(5):63-70. http://www.ncbi.nlm.nih.gov/pubmed/29114196%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5669513.
5. Zerjav Tansek M, Groselj U, Angelkova N, et al. Phenylketonuria screening and management in southeastern Europe - Survey results from 11 countries. *Orphanet J Rare Dis*. 2015;10(1):1-7. doi:10.1186/s13023-015-0283-0
6. Chiril Boiciuc, Natalia Uşurelu, Mihail Strătilă VS. Fenilcetonuria În Republica Moldova. Diagnosticul prin screening neonatal şi analiza molecular-genetică. *A VI- a Conferinţă Internaţională - Zilele Neonatol Mold*. 2013;Văratec:146-153.

7. Chiril Boiciuc, Natalia Uşurelu VS. Analiza molecular-genetică a genei PAH la pacienţii cu PKU din Republica Moldova. *Bul. Acad. Ştiinţe a Mold. Ştiinţe Medicale*. 2014:227-232.
8. Anna A, Monika G. Splicing mutations in human genetic disorders : examples , detection , and confirmation. 2018:253-268.
9. K. Boiciuc, N. Uşurelu VS. *Metode de diagnostic clinic și de laborator în genetica medicală*; 2019.
10. Chiril Boiciuc VS. Secvenţierea ADN-ULUI. *METOD DIAGNOSTIC Clin. și lab. în genet. MEDICALĂ*. 2019:55-63.
11. Dunnen JT Den, Dalgleish R, Maglott DR, et al. HGVS Recommendations for the Description of Sequence Variants : 2016 Update. 2016. doi:10.1002/humu.22981
12. Richards S, Aziz N, Bale S, Bick D, Das S. ACMG Standards and Guidelines 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. 2015;17(5):405-424. doi:10.1038/gim.2015.30
13. Brunak S, Engelbrecht J. Prediction of Human mRNA Donor and Acceptor Sites from the DNA Sequence. 1991:49-65.