

Molecular genetics and diagnosis of phenylketonuria: state of the art

Expert Rev. Mol. Diagn. 14(6), 655–671 (2014)

Nenad Blau^{*1,2},
Nan Shen¹ and
Carla Carducci³

¹Division of Inborn Metabolic Diseases,
University Children's Hospital, Im
Neuenheimer Feld 430, 69120
Heidelberg, Germany

²Division of Metabolism, University
Children's Hospital, Zürich, Switzerland

³Laboratory of Genetics and Metabolic
Diseases, Division of Experimental
Medicine, University Sapienza, Rome,
Italy

*Author for correspondence:
nenad.blau@med.uni-heidelberg.de

Detection of individuals with phenylketonuria (PKU), an autosomal recessively inherited disorder in phenylalanine degradation, is straightforward and efficient due to newborn screening programs. A recent introduction of the pharmacological treatment option emerged rapid development of molecular testing. However, variants responsible for PKU do not all suppress enzyme activity to the same extent. A spectrum of over 850 variants, gives rise to a continuum of hyperphenylalaninemia from very mild, requiring no intervention, to severe classical PKU, requiring urgent intervention. Locus-specific and genotypes database are today an invaluable resource of information for more efficient classification and management of patients. The high-tech molecular methods allow patients' genotype to be obtained in a few days, especially if each laboratory develops a panel for the most frequent variants in the corresponding population.

KEYWORDS: BH4 • hyperphenylalaninemia • kuvan • PKU • tetrahydrobiopterin

History of phenylketonuria

Phenylketonuria (PKU) was the first known metabolic disorder recognized to cause mental retardation [1]. It was also the first genetic disorder for which a treatment could prevent its most devastating effects. The history of PKU includes a continuous understanding of the mechanisms of the enzyme deficiency and a continuous journey toward better treatment [2].

One of the biochemical hallmarks noted first was a peculiar musty odor of patient's urine. It was Norwegian physician, Asbjørn Følling who tested their urine with a ferric chloride solution to detect the presence of ketones. This test yielded a reaction (a transient dark green color) never previously described. Følling identified the substance producing the color changes as phenylpyruvic acid. In this way, in 1934, he described a biochemical disorder associated with mental retardation, the first such description [3]. The name, 'phenylketonuria', was given to this disorder by Lionel Penrose, the first to recognize the importance of Følling's discovery [4]. Later, George Jervis, the pioneer of research into PKU in the USA, identified a defect in the function of the phenylalanine hydroxylase (PAH) enzyme as the cause of the

hyperphenylalaninemia (HPA) that has characterized PKU [5].

The first successful dietary intervention in the management of PKU occurred in the 1950s. Horst Bickel, a young physician from Germany, was at that time in training at the Children's Hospital in Birmingham. He saw a girl demonstrated with what we now regard as classic symptoms of untreated PKU: mental retardation without development of language or interest in her surroundings, abnormal gait, spastic reflexes, eczema and aberrant behaviors. Building on the work of Følling and others, Bickel identified PKU in the girl and he and a colleague from London, Louis Woolf, prepared the first phenylalanine (Phe)-free protein substitute [6]. The child demonstrated dramatic improvement in behavior, mood and motor development, to crawl and stand. This showed that Phe restriction held the key to improving outcome in PKU. Soon the Phe formula became standard treatment for this disorder, with greatest benefits when initiated in the newborn period.

The development of the first routine screening test for PKU in the early 1960s by Robert Guthrie, based on measurement of Phe concentrations in blood spots, dramatically changed this equation [7]. This test made

possible neonatal screening for PKU and all other forms of HPA, allowing the Phe-restricted diet to begin in time to prevent the severe consequences of PKU. Work in the 1980s and 1990s localized the human *PAH* gene to chromosome 12 [8]. Additional research during the intervening three decades has identified more than 850 variants of the *PAH* gene [9–11].

Clinical presentation, pathophysiology, treatment & outcome

The clinical presentation of children with PKU is unremarkable in the period immediately after birth. Untreated children with classic PKU, however, develop severe intellectual disability, seizures, ataxia, motor deficits and behavioral problems, and, in many cases, features of autism. From an age of about 6 months, these children exhibit developmental problems, which may be accompanied by aberrant behaviors including self-harm, aggression, impulsivity and psychosis. Young children with a late diagnosis of PKU display marked and rapid benefit from dietary treatment. They may commence walking and talking and demonstrate improved intellectual functioning and behavior. However, most retain a clear degree of intellectual disability and almost all display significant learning disabilities, irrespective of some improvement in IQ. The degree of intellectual disability depends on the age at which treatment is started. Improved cognitive performance following Phe restriction in adults with previously untreated PKU may have been facilitated by decreased aggression, psychosis or self-harm. These maladaptive behaviors were refractory to medications and behavioral programs, but diminished with long-term dietary Phe control [12–16].

PKU pathology is almost completely restricted to the brain. The etiology of cognitive problems arising as a result of HPA is unclear, and several mechanisms may be involved at various levels. There is general agreement on the importance of the blood–brain barrier. Phe competes with other large neutral amino acids (LNAA) for a carrier protein responsible for their uptake into the CNS. Thus, HPA simultaneously increases the brain Phe level and reduces levels in the brain of other amino acids essential for protein synthesis and neurotransmitter production. Accordingly, a reduced level of LNAA other than Phe should be considered in its own right as a contributor to the pathophysiology of neurological impairment in PKU. Various additional mechanisms that may be involved in the brain damage: altered myelin metabolism (high concentrations of Phe and/or reduced availability of other LNAA may inhibit the development of myelin, leading to lack of myelin formation in untreated patients, and to edema within myelin in early-treated patients with PKU), impaired dopamine and serotonin synthesis (high brain Phe concentration inhibit the rate-limiting enzymes in the biosynthesis of biogenic amines; tyrosine and tryptophan hydroxylases), impairment of glutamatergic synaptic transmission and the activity of key enzymes, such as pyruvate kinase, tyrosine and tryptophan hydroxylase and HMG-CoA reductase (all due to high brain Phe concentrations) [17–19].

Optimal psychosocial and neurocognitive outcome should be the aim of diagnostic process and treatment [20]. Thus, treatment should start as soon as possible after birth and continue throughout life, according to national guidelines. Diet is the basis of treatment for patients with PKU with elevated blood Phe concentrations [21]. It consists of a limited and controlled amount of natural protein derived from food sources to provide essential Phe requirements. The majority of nutrient requirements are met by a Phe-free source of L-amino acids that are commonly supplemented with added micronutrients. In most countries with PKU screening programs, dietary treatment is initiated on confirmation of a positive screening test and at least within 21 days. The overall aim of treatment for PKU should be to ‘achieve an optimal cognitive and psychological development and wellbeing of the child’. In general, the usual treatment goal is to achieve maintenance of blood Phe concentrations between 120 and 360 $\mu\text{mol/l}$, although there is variation between clinics and across countries. There is little agreement about the blood Phe target concentrations for adolescents and adults with PKU and there is no international consensus on the need for long-term treatment. Generally, blood Phe targets become less stringent with increasing age. Typically, the target blood Phe range becomes higher at 4 or 8 years of age, with further relaxations between 10 and 16 years, and adults, with target Phe concentrations varying between countries [1,22,23].

Classification, nomenclature & phenotypes

PKU requiring therapeutic intervention to prevent neurological damage is usually diagnosed on the basis of blood Phe $>360 \mu\text{mol/l}$ (see below). However, the classification of HPA and PKU is not simple [24]. Previously, it could be based on highest blood Phe levels before introducing treatment (metabolic phenotype). Patients with blood Phe levels of 120–600 $\mu\text{mol/l}$ were classified as mild HPA (MHP), blood Phe levels of 600–1200 $\mu\text{mol/l}$ were classified as mild PKU and levels above 1200 $\mu\text{mol/l}$ as classic PKU. However, this can only be applied when Phe reaches its full potential biological value due to a longer period without treatment. Classification is not accurate when the Phe level used for classification is obtained from newborn screening and may not have had time to reach the highest level. Thus, classification must also consider the tolerance for dietary Phe [25].

Distribution of metabolic phenotypes is different in different populations and depends on the genotype. The genotype determines protein dysfunction, residual PAH activity and consequently the phenotype. It has been reported that classic PKU is predominant phenotype in Eastern European countries, where severe variants with almost no residual enzyme activity (e.g., **c.1222C>T**) are more frequent, while in the southern part of Europe mild HPA is more common, due to frequent mild variants with a substantial residual PAH activity (e.g., c.1169A>G). Based on the data from 7405 patients from all world regions tabulated in the BIOPKU database [26], half of them (54.8%) are classic PKU, 27.4% mild PKU and 17.8% MHP [11].

Regardless of the metabolic phenotype, patients may develop different severity of the disease (clinical phenotype). Individuals with the same metabolic phenotype or even same genotype may have different cognitive impairment. Some adult patients with classic PKU and high blood Phe levels remain unaffected on free diet, while others suffer from a severe mental disability. This phenomenon seems to be explained by different Phe concentration in the blood and brain. Blood–brain barrier and transport of LNAA may play a central role in determining the clinical phenotype [27]. Additional possible players may be gene modifiers [28].

Tetrahydrobiopterin responsiveness

The phenomenon of tetrahydrobiopterin (BH₄)-sensitive PKU was initially described in Japanese patients by Kure *et al.* [29] and confirmed in two retrospective studies with 1730 and 557 patients with PAH deficiency [30,31]. *In vivo* functional test confirmed impaired Phe oxidation in the liver [32]. In these patients, oral administration of 10–20 mg BH₄ per kg body weight reduced blood Phe concentrations significantly [33]. The proportions of patients responding strongly to BH₄ decreased as the severity of HPA increased. Thus, a standardized BH₄ loading test is required to optimize the identification of patients who might benefit from pharmacological treatment with BH₄. Several tests have been proposed, using different dosages of BH₄ (usually 10 or 20 mg/kg), repeated dosages of BH₄ or co-administration of BH₄ with Phe [33]. The definition of a clinically significant response to BH₄ treatment, in terms of a reduction in blood Phe of at least 30%, is generally accepted for use within these tests. Larger response rates were consistently observed in patients with milder variants of HPA or PKU. The efficacy and safety of this relatively new treatment for PKU was documented in a number of well-designed studies [34,35]. Overall, considering PKU of any severity, a response rate to BH₄ load of about 30–40% was observed in the studies that employed the usual cut-off for efficacy of a reduction in blood Phe of at least 30%. In the future, this methods used in conjunction with genotyping (see below) might permit a more robust quantification of the degree of BH₄ responsiveness in an individual. Currently, the most used modality in Europe is a 48-h challenge with 20 mg BH₄ per kg body weight and day [36,37]. This loading test essentially consists of a single-dose, single-day evaluation over 2 days. All of the proposed loading tests involved reassessment of the diet, measurement of Phe tolerance and titration of BH₄ in responders to treatment. However, it should be stressed that with regard to the long-term efficiency, the initial short-time BH₄ test may detect some patients, in whom the reduction of blood Phe levels is not sufficient to replace the classical dietary treatment. On the contrary, some patients may need a longer period of time for testing and are thus characterized as late or slow responder. A study by Shintaku *et al.* describes different testing modalities over several weeks and with different dosages of BH₄ (10–20 mg/kg/d) [38].

To some extent, genotype predicts the degree to which blood Phe will respond to treatment with BH₄ in patients with PAH-

deficient PKU [10,39,40]. The variants in the *PAH* gene that gave rise to BH₄-responsive PKU were examined in several studies, using data from the BIOPKU database (see also below). Initially, it was believed that a single BH₄-responsive variant might be informative enough [41]. Although, while the genotype is partially predictive of BH₄ responsiveness, it is not currently possible to predict this phenomenon with 100% confidence in an individual patient based on genotyping alone [42,43].

Epidemiology

The prevalence of PKU is highest in Caucasian or East Asian subjects, at about 1:10,000–15,000 births. However, the number of cases of PKU is not distributed evenly: for example, persistent HPA occurs in about 1:6500 births in Turkey and Northern Ireland. The prevalence in Latin America varies from about 1:25,000 to about 1:50,000 births, with a generally lower prevalence in the north. Reported prevalence rates in Asia vary from about 1:11,500 to 1:100,000 in regions of China, less than 1:200,000 in Thailand and about 1:70,000 in Japan. Spain differs from other European countries in having a relatively high incidence of mild elevations of blood Phe arising from partial inactivation of PAH. The prevalence of PKU in subjects of African or South Asian descent may be lower than the prevalence in Caucasian populations [1].

Phe hydroxylating system

PAH is a mixed function oxidase that catalyzes hydroxylation of Phe to tyrosine, the rate-limiting step in Phe catabolism [44]. The reaction is dependent on BH₄ as cofactor, molecular oxygen and iron. Additional components coupled with this system that allow normal function are pterin-4 α -carbinolamine dehydratase, which eliminates water from the oxidized pterin, and dihydropteridine reductase, which regenerates BH₄ by using NADH (FIGURE 1) [45].

PAH is present in the liver due to its function of metabolizing dietary Phe. The human organism is thought not to have measurable PAH activity in kidney. However, in mice kidney PAH activity is about 30% of the liver enzyme [46].

Human PAH is a cytosolic protein and exists in solution as a pH-dependent equilibrium between functional tetramers and dimers [47]. Like the other two aromatic amino acid hydroxylases tyrosine hydroxylase and tryptophan hydroxylase, PAH consists of three domains: the regulatory domain (residues 1–142), the catalytic domain (residues 143–410) and the tetramerization domain (residues 411–452) (FIGURE 2).

PAH activity is highly regulated by reversible phosphorylation, substrate activation and cofactor inhibition [48,49]. Within the first 30 residues of PAH, there is an autoregulatory sequence that includes the Ser16 (substrate for cAMP-dependent protein kinase) and residues essential for positive cooperativity by the substrate and the inhibitory effect by BH₄ [50].

Phe hydroxylase gene

In 1984, Lidsky *et al.* assigned the *PAH* gene to human chromosome 12 (12q22-12q24.2) and in 1986 the molecular structure of

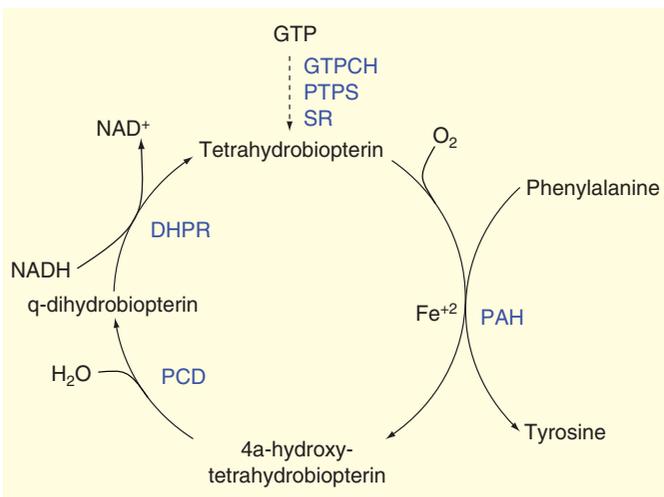


Figure 1. The phenylalanine hydroxylating system showing hydroxylation of phenylalanine to tyrosine in the presence of molecular oxygen, iron and co-factor/co-substrate tetrahydrobiopterin. During the hydroxylation event, tetrahydrobiopterin is oxidized to a hydroxyl-cabonolamine intermediate, which is consecutively reduced/recycled back to tetrahydrobiopterin by enzymes carbinolamine-4a-dehydratase and dihydropteridine reductase.

DHPR: Dihydropteridine reductase; GTPCH: GTP cyclohydrolase I; PAH: Phenylalanine hydroxylase; PCD: Pterin-4a-carbinolamine dehydratase; PTPS: 6-pyruvoyl-tetrahydropterin synthase; SR: Sepiapterin reductase.

the gene was finally determined [8,51]. DNA size is 79.28 kb (NG_008690.1, NT_029419.13, NC_018923.2, NC_000012.12), and it is composed of 13 exons. The exons are not equally distributed: in the second moiety of the gene the exons are compact, whereas those in the first moiety are separated by large introns. The

transcription occurs on the reverse strand and only one isoform has been identified. The mature mRNA is of 2680 bp (NM_000277.1) [52], while the CDS is 1359 bp [53]. The 5' untranslated region region was characterized in 1992 by Konecki *et al.* [54]. There is a cap site for the transcription 154 nucleotides upstream of the first ATG codon. The promoter region is characterized by the lack of a proximal TATA box, but there is the presence of GC boxes, CACCC boxes, CCAAT boxes, activator protein 2 sites, partial glucocorticoid response elements and partial cyclic AMP response elements consensus sequence. This suggests that the human *PAH* gene has a TATA-less promoter regulated by multiple transcription factors. Three polyadenylation signals [AATAAA] in exon 13 are annotated on the gDNA sequence. The *PAH* gene ID is 5053 [55].

Locus-specific database PAHvdb

The locus-specific variations database *PAHvdb* [56] is an open access resource for variations reported in the *PAH* gene. It is structured and maintained according to the nomenclature standards recommendations by the human genome variation society [57]. In addition to the core information (e.g., *PAH* identifier, nucleotide aberration, protein variant, gene region, protein domain, variation type, *in vitro* enzyme activity, etc.), information on protein damage and pathogenic effect of the variations are calculated using prediction tool algorithms (e.g., FoldX, Polyphen 2, SIFT and SNPs3D). Allelic phenotype value (APV) was calculated for most variant. The 3D atomic environment of each missense variant was visualized using the interactive iSee concept [11]. A total of 851 variants were tabulated in the *PAHvdb* (as of May 2014). Summary of variants is shown in TABLE 1 & FIGURE 3. Most of the variants (61.2%) were located in the catalytic domains, followed by the regulatory (16.8%) and oligomerization (5.2%) domains. The rest of variants were located in non-coding regions of

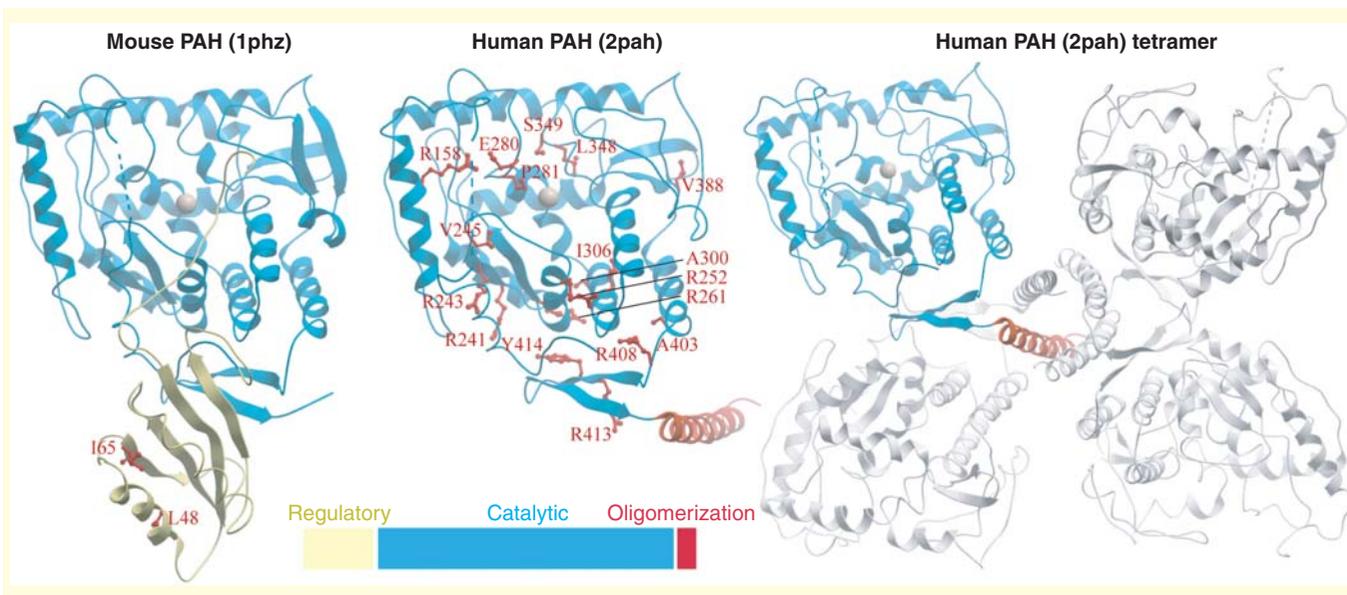


Figure 2. The 3D structure of the phenylalanine hydroxylase monomeric subunit with locations of the most common missense variants (for a complete list see TABLE 3).

Original image provided by Wyatt W. Yue, SGC, Oxford, UK.

Table 1. Summary of phenylalanine hydroxylase gene variants tabulated in the PAHvdb database.

Gene region	Protein domain				Variant type						Total	
	Regulatory	Catalytic	Oligomerization	None	Missense	Nonsense	Splice	Silent	Deletion	Insertion		Other [†]
5'UTR				8				5	3			8
Exon 1	17			1	8	1		1	5	3		18
IVS 1				6			6					6
Exon 2	33			11	22	1		1	7	2		33
Exon 3	77			7	55				17	3	2	77
IVS 3				7			7					7
Exon 4	15	5		16	16	2		2	2			20
IVS 4				16			14		2			16
Exon 5		46		12	35	3		4	4		4	46
IVS 5				12			12					12
Exon 6		116		5	83	15		15	2	1	1	116
IVS 6				5			5					5
Exon 7		117		11	98	4		12	1	2	2	117
IVS 7				11			11					11
Exon 8		37		9	29	3		1	3	1	1	37
IVS 8				9			9					9
Exon 9		37		13	28			8		1	1	37
IVS 9				13			13					13
Exon 10		62		17	46	7		8	1			62
IVS 10				17			17					17
Exon 11		82		13	57	7		1	14	3		82
IVS 11				13			13					13
Exon 12		17	39	10	42	5		1	7		1	56
IVS 12				10			10					10
Exon 13			5	3	2	1		1	1	1		5
3'UTR				3			3					3
Total	142	519	44	142	521	49	127	13	108	16	13	847

[†]Indels, large deletions.
Data taken from [56].

the gene. Variants were spread out through the entire *PAH* gene, but most affected regions were exons 6, 7 and 11, with 13.7, 13.8 and 9.7% of variants, respectively (FIGURE 3). Out of 851 gene variations, 61.5% were missense variants, 15.0% splice variants, 12.8% deletions, 5.8% nonsense variants, 1.9% insertions and 3.0% included indels, large deletions and silent variations. *PAH*db is linked with the genotypes database BIOPKU (see below), OMIM [58], the human gene mutation database [59] and disease-centered central variant databases [60].

Diagnosis of HPAs

HPAs and PKU diagnosis is achieved soon after birth by neonatal screening in most developed countries [25]. In the countries in which expanded newborn screening has been adopted, PKU is diagnosed by Phe and Tyr evaluation in neonatal dried blood spot (DBS) by using tandem mass spectrometry. Virtually, 100% of PKU subjects should be identified. The risk of false negative has become more pressing with early postnatal discharge from hospital for mothers and term infants. In such cases, Phe/Tyr ratio evaluation increase sensitivity and predictive positive value of the test, avoiding the false-negative and lowering false-positive risk. Other methods used for PKU neonatal screening, although still utilized, do not allow the simultaneous evaluation of Tyr in the neonatal DBS. Bacterial inhibition assay (Guthrie test) is a simple, inexpensive and reliable test used for many decades; however, it is a manual, semi-quantitative test and is being replaced by other methods in quite all screening laboratories. Some laboratories use a fluorimetric test that is quantitative, automated and reliable.

Immediately after the first positive test, it is necessary to confirm the positive Phe value. The confirmation methods, as in all screening programs, should show a higher specificity to exclude false positives. In this case, Phe and Tyr should be tested in plasma by using tandem mass spectrometry or even better with chromatographic methods (ionic-exchange or HPLC).

Because of genetic heterogeneity, once the first screening and confirmation tests show blood Phe concentration >120 $\mu\text{mol/l}$ and/or Phe/Tyr ratio >2 (internationally accepted cut-offs), it is necessary to exclude a defect in BH₄ pathway genes [25]. In about 98% of cases, Phe increase is due to a PAH defect and in 2% due to defects in BH₄ synthesis and regenerating genes. Although BH₄ disorders are rare, disease progression is usually much more severe than in PKU patients and affected newborns need immediate pharmacological intervention. Therefore, it is mandatory to exclude BH₄ deficiency in all newborns with even mild HPA by pterin assessment in DBS or urine and dihydropteridine reductase activity.

PAH molecular analysis

Rat *PAH* cDNA was cloned in 1982 by Roboson *et al.* [61]. It was then used to screen a human liver cDNA library to obtain human *PAH* cDNA. Human *PAH* cDNA, engineered into a bacterial expression vector, could be grown into *Escherichia coli* and used in a hybridization assay. In particular, at human *PAH* locus eight polymorphic sites recognized by seven restriction enzymes were

identified. Genomic DNA was digested by the 7 enzymes and, using Southern blot technique, hybridized to a full-length human *PAH* cDNA clone. The resulting pattern showing restriction fragment length polymorphisms gave rise to a haplotype that segregated in a Mendelian fashion and concordantly with the disease state in various PKU kindreds. In Caucasians, this haplotype resulted informative in 87.5% of population [62].

Following the discovery of gene structure, the methods used for the study of the molecular bases of PKU go hand in hand with technological innovations. In 1977, Frederick Sanger developed a rapid method for DNA sequencing that allowed the first results in gene sequencing. *PAH* gene sequencing analysis of a mutant clone allowed to identify the first variant c.1315+1G>A causing disease in Danish population [51]. Soon after a second variant c.1222C>T (p.Arg408Trp) was identified by the same method [63].

Using DNA amplification by PCR and specific hybridization oligonucleotides labeled by radioisotopes, DiLella *et al.* [64] described a rapid method for the analysis of the two known mutant alleles, frequent among Northern Europeans: c.1315+1G>A and c.1222C>T (p.Arg408Trp). This results seemed decisive for *PAH* genotyping since in the population analyzed these two variants reached a detection rate of 60% and the allelic variability of the *PAH* gene among different populations was not yet known.

Template preparation based on PCR amplification allowed avoiding awkward cloning step and the use of cycle sequencing procedure smooth the way to studies of PKU genetics on different populations. In the following years, many new variants were reported and the allelic heterogeneity of the disease became evident [65]. However, the detection system was still based on use of radioactive isotopes that makes this procedure time consuming and difficult in handling. This problem was overcome by development of exon scanning techniques. One of the still most widely used for *PAH* variant analysis is denaturing gradient gel electrophoresis. It is a particular type of gel electrophoresis that uses an increasing concentration of denaturing chemicals at a constant heat (about 60°C) to strain DNA to unwind (or 'to melt'). During migration, DNA reaches the concentration of denaturing reagents at which it melts. At this concentration, the migration along the gel is so slowed down that it can be considered absent. Any variation in DNA sequence will result in different melting temperatures, and consequently in different migration in the gel. It is possible to enhance the difference between wild-type and mutant allowing heteroduplex formation. In fact, hybrids formed between different alleles, also for a single mismatch (heteroduplex), migrate slower than respective homoduplex, creating during electrophoresis a characteristic pattern (FIGURE 4A). Heteroduplex can also be generated by artificially mixing target DNA to a wild-type DNA. Using this approach, only exons carrying a sequence variant are submitted to sequencing analysis. A criticality is the need of a 40 nt GC-clamp at 5'-end of one primer of the couple used for amplification to generate a unique melting domain in the amplicon to be analyzed. This exon scanning technique showed a very high analytical sensitivity and specificity (>96%) [66].

Expert Review of Molecular Diagnostics Downloaded from informahealthcare.com by 212.35.22.196 on 06/19/14
For personal use only.

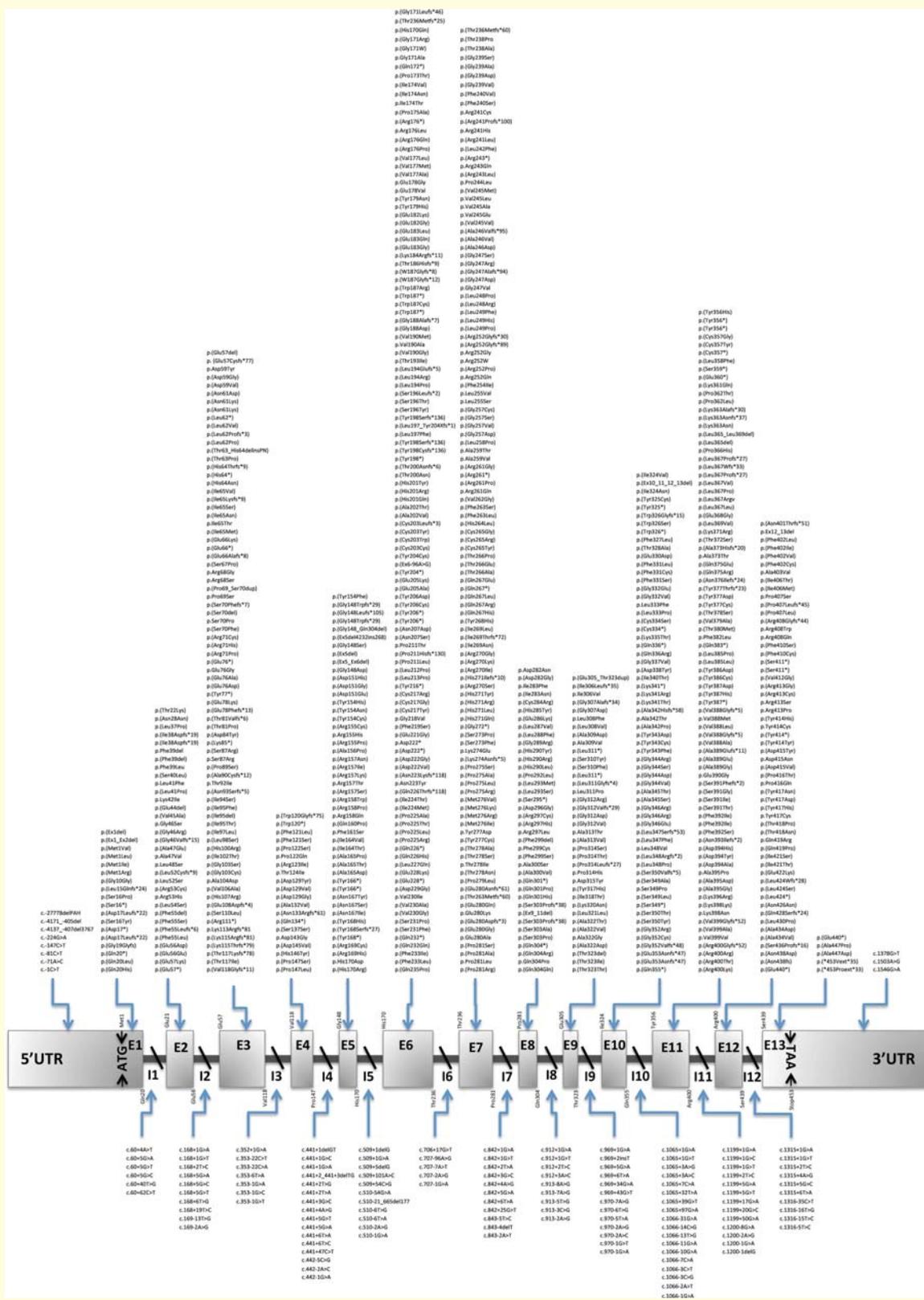


Figure 3. Physical structure of the phenylalanine hydroxylase gene with variants listed in the PAHdb locus-specific database. Data taken from [56].

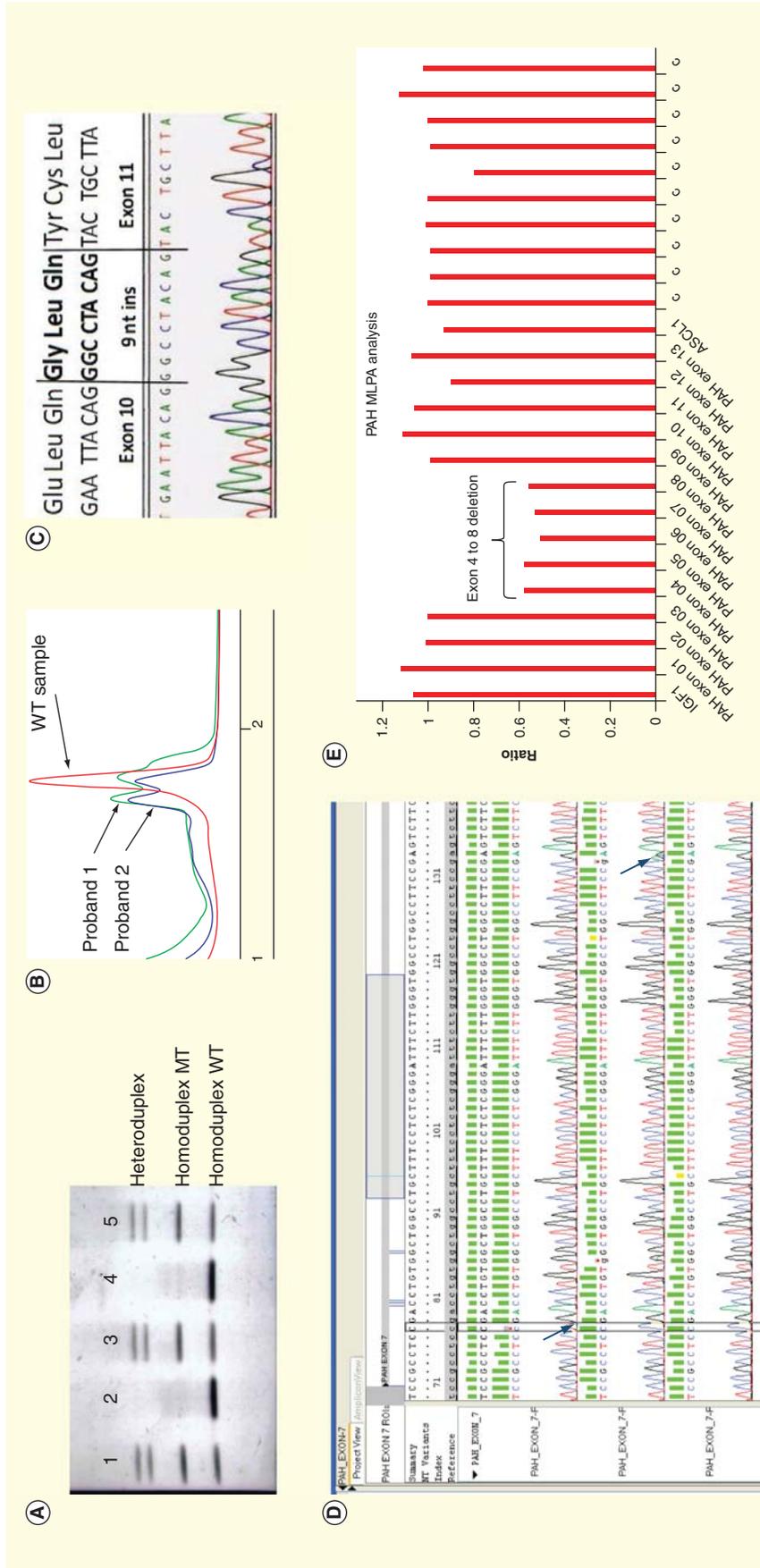


Figure 4. Molecular analysis in patients with phenylketonuria. (A) Phenylalanine hydroxylase exon 11 analysis by denaturing gradient gel electrophoresis of a family carrying c.1066-11G>A variant (lanes 1 = father, 3 = proband 1, 5 = proband 2; lanes 2 and 4 = mother and sibling wild type); (B) denaturing high-pressure liquid chromatography analysis of two samples carrying p.Arg252Trp (proband 1 and 2); (C) analysis of mRNA illegitimate transcript of a proband carrying c.1066-11G>A variant. The variant causes a 9 intronic nt retention with an insertion of 3 amino acids (Gly, Leu, Gln); (D) results of exon 7 sequencing analysis interpretation by alignment software (SeqScape, Applied Biosystems, Foster City, CA, USA). The first raw represents a sample carrying p.(Arg243*) in heterozygosis (arrow), the second raw shows a sample carrying p.Arg261Gln in heterozygosis, the third raw is a wild-type sample; (E) multiple ligations probe amplification analysis of a proband carrying a heterozygous deletion encompassing exons 4 to 8.

Another exon scanning technique, very promising and with a high grade of automation, is denaturing high-pressure liquid chromatography. Denaturing high-pressure liquid chromatography is a technique for separation and analysis of DNA fragments with different length and/or base compositions. It is based on a reversed-phase chromatographic system in which the column is maintained at elevated temperature that allows a partial DNA denaturation. In these conditions, DNA homoduplex and heteroduplex show a different retention time allowing identifying fragments carrying a sequence variant (FIGURE 4B). This technique has a very high sensitivity and specificity (near to 100%) to detect known variants but less to screen for unknown, since the analysis conditions are highly specific for each single nucleotide variation. This method adapted more to variant screening than to exon scanning [67]. A disadvantage of exon scanning technique is the high rate of known polymorphisms found in *PAH* gene, that constrain to carry out sequencing analysis on a high number of fragments.

The advent of multicapillary automated genetic analyzers and Big Dye Terminators fluorescently labeled used in a cycle sequencing reaction made the sequencing analysis more rapid, high-throughput and less expensive. In this way, the exon scanning approach could be avoided. The analytical sensitivity and specificity of this method is very high (>99%), however, the diagnostic sensitivity is lower (FIGURE 4D). The traditional sequencing approach evaluates genes exon by exon including intron-exon boundaries to detect splicing variants, but excluding intronic regions from the analysis. Furthermore, as for all PCR-based methods, it is not able to determine the intra-gene deletions/duplications. Several studies have demonstrated that intra-genic deletions/duplications are a relative frequent event in *PAH* gene. Patients with none or only one variant should be screened for deletions/duplications using multiple ligations probe amplification technique [68]. This technique is a semi-quantitative exon-by-exon amplification, in which reaction condition and normalization to several standards allows determining if each exon is present, absent or 50% reduced (FIGURE 4E). Several reports described an increase in diagnostic sensitivity using this approach. Nevertheless, the diagnostic sensitivity, although increased, does not reach 100% [69].

More than 850 variants have been reported until now in *PAH* gene (allelic heterogeneity), and it is ever possible to detect very rare or new variants. In such a case, it is necessary to assign a pathologic role to the variant. Patient's enzymatic activity cannot be used since the enzyme is at hepatic expression and the *in vitro* expression of mutated protein is a time-consuming and expensive procedure. The locus-specific database can be very useful: if a variant is identified for the first time it is useful to know if in another laboratory worldwide it was already observed in association with the disease. Simultaneously, today we have a series of bioinformatics tools that help us in finding new variant interpretation [10,46,70,71].

The *PAH* mRNA analysis could be useful to exclude splicing defects due to intronic variants, that are not detectable using sequencing testing. In the past, studies have been conducted to

analyze *PAH* mRNA, starting from the observation that tissue-specific mRNAs are transcribed and processed at basal level in adult and embryonic tissues [72]. Illegitimately transcribed *PAH* mRNA was amplified using the PCR from both fibroblasts and Epstein–Barr virus-transformed lymphocytes (FIGURE 4C); however, these studies are long and complex, and their usefulness has mainly research purposes [73,74].

Utility of genetic testing

Molecular *PAH* analysis is not necessary for the diagnosis. However, it can be very useful also in neonatal period. The up-to-date molecular methods allow obtaining patients' genotype in few days, especially if each laboratory develops a rapid test for the most frequent variants in the belonging population. This information, obtained early after birth, can help clinicians in differential diagnosis excluding or confirming a defect in BH₄ synthesis or regenerating genes. Moreover, there is, although with few exceptions, a good genotype–phenotype correlation (see section 'Genotype–phenotype correlation') that can be useful to clinicians for phenotypic overview. This aspect has become more important recently with early sampling due to early postnatal discharge. With this timing, the diagnosis is made very early and the therapy established before Phe concentration has reached its maximum. This shows that the patient is not assigned to correct phenotypic category, considering also the fact that the specific test (Phe tolerance) can be performed at 5 years of age.

PKU genotyping can be useful for therapy as well. There is a strong relationship between variants and BH₄ responsiveness (see section 'Tetrahydrobiopterin responsiveness') and many research studies on variant-based therapy (see below) have been conducted.

For genetic counseling in PKU families it is very important to assess parent's plasma Phe values. In case of one parent's asymptomatic HPA, the presence of two mutated alleles, which can segregate differently, could give rise to different phenotype in following pregnancies (from HPA to severe PKU). This event explains some of the intra-familial phenotypic inconsistencies observed. Finally, genetic characterization of PKU families can allow prenatal diagnosis in following pregnancies. Nevertheless, being the disease treatable, this is not a frequent request.

Genotypes database BIOPKU

BIOPKU is a database of PKU patients (all phenotypes) with a full genotype [26], linked to the locus-specific database PAHvdb (see above). A total of 7405 patients with 1918 different genotypes had been tabulated in BIOPKU as of May 2014. In addition to the genotype information, BIOPKU includes information on the patient's phenotype, responsiveness to BH₄, blood Phe concentrations in the NBS and tolerance to dietary Phe intake (if applicable). TABLE 2 shows the most common genotypes and TABLE 3 shows the most common alleles.

Phenotype was reported for 6513 (88.0%) patients. The smallest group was MHP comprising 1161 patients (15.7%), followed by mild PKU (24.1%) and classic PKU (48.2%). These numbers

Table 2. The most frequent genotypes in the BIOPKU database occurring in $\geq 0.4\%$ of all patients.

Genotype	Protein variation	Genotype frequency (%)
c.[1222C>T];[1222C>T]	p.[Arg408Trp];[Arg408Trp]	10.8
c.[1066-11G>A];[1066-11G>A]	p.[?];[?]	2.2
c.[782G>A];[782G>A]	p.[Arg261Gln];[Arg261Gln]	1.7
c.[473G>A];[1222C>T]	p.[Arg158Gln];[Arg408Trp]	1.6
c.[1222C>T];[1315+1G>A]	p.[Arg408Trp];[?]	1.5
c.[782G>A];[1222C>T]	p.[Arg261Gln];[Arg408Trp]	1.4
c.[1208C>T];[1222C>T]	p.[Ala403Val];[Arg408Trp]	1.2
c.[1241A>G];[1222C>T]	p.[Tyr414Cys]; [Arg408Trp]	1.1
c.[842C>T];[1222C>T]	p.[Pro281Leu];[Arg408Trp]	1.1
c.[842C>T];[842C>T]	p.[Pro281Leu];[Pro281Leu]	1.0
c.[143T>C];[1222C>T]	p.[Leu48Ser];[Arg408Trp]	0.9
c.[1066-11G>A];[1222C>T]	p.[?];[Arg408Trp]	0.8
c.[143T>C];[143T>C]	p.[Leu48Ser];[Leu48Ser]	0.8
c.[898G>T];[1222C>T]	p.[Ala300Ser];[Arg408Trp]	0.8
c.[1066-11G>A];[1222C>T]	p.[?];[Arg408Trp]	0.7
c.[782G>A];[1066-11G>A]	p.[Arg261Gln];[?]	0.7
c.[1241A>G];[1315+1G>A]	p.[Tyr414Cys];[?]	0.6
c.[1169A>G];[1222C>T]	p.[Glu390Gly];[Arg408Trp]	0.6
c.[754C>T];[1222C>T]	p.[Arg252Trp];[Arg408Trp]	0.6
c.[1315+1G>A];[1315+1G>A]	p.[?];[?]	0.6
c.[143T>C];[782G>A]	p.[Leu48Ser];[Arg261Gln]	0.5
c.[916A>G];[1222C>T]	p.[Ile306Val];[Arg408Trp]	0.5
c.[728G>A];[728G>A]	p.[Arg243Gln];[Arg243Gln]	0.5
c.[838G>A];[838G>A]	p.[Glu280Lys];[Glu280Lys]	0.5
c.[194T>C];[1222C>T]	p.[Ile65Thr];[Arg408Trp]	0.5
c.[1042C>G];[1222C>T]	p.[Leu348Val];[Arg408Trp]	0.4
c.[754C>T];[754C>T]	p.[Arg252Trp];[Arg252Trp]	0.4
c.[782G>A];[1315+1G>A]	p.[Arg261Gln];[?]	0.4
c.[165delT];[1222C>T]	p.[Phe55Leufs*6];[Arg408Trp]	0.4
c.[1241A>G];[1241A>G]	p.[Tyr414Cys];[Tyr414Cys]	0.4
c.[194T>C];[1066-11G>A]	p.[Ile65Thr];[?]	0.4
c.[1241A>G];[1066-11G>A]	p.[Tyr414Cys];[?]	0.4
c.[143T>C];[1066-11G>A]	p.[Leu48Ser];[?]	0.4
c.[473G>A];[473G>A]	p.[Arg158Gln];[Arg158Gln]	0.4
c.[898G>T];[1066-11G>A]	p.[Ala300Ser];[?]	0.4

Data taken from [26].

Table 3. The most frequent alleles in the BIOPKU database occurring in $\geq 0.5\%$ of all patients.

cDNA aberration	Protein aberration	Gene location	Protein domain	Allele frequency (%)
c.1222C>T	p.Arg408Trp	Exon 12	Catalytic	22.55
c.1066-11G>A	p.(?)	IVS 10		6.07
c.782G>A	p.Arg261Gln	Exon 7	Catalytic	5.60
c.1315+1G>A	p.(?)	IVS 12		3.64
c.842C>T	p.Pro281Leu	Exon 7	Catalytic	3.42
c.473G>A	p.Arg158Gln	Exon 5	Catalytic	3.15
c.143T>C	p.Leu48Ser	Exon 2	Regulatory	3.07
c.1241A>G	p.Tyr414Cys	Exon 12	Oligomerization	3.04
c.1208C>T	p.Ala403Val	Exon 12	Catalytic	2.48
c.194T>C	p.Ile65Thr	Exon 3	Regulatory	2.39
c.728G>A	p.Arg243Gln	Exon 7	Catalytic	2.35
c.898G>T	p.Ala300Ser	Exon 8	Catalytic	1.57
c.754C>T	p.Arg252Trp	Exon 7	Catalytic	1.52
c.1162G>A	p.Val388Met	Exon 11	Catalytic	1.50
c.727C>T	p.(Arg243*)	Exon 7	Catalytic	1.18
c.721C>T	p.Arg241Cys	Exon 7	Catalytic	1.09
c.838G>A	p.Glu280Lys	Exon 7	Catalytic	1.08
c.165delT	p.(Phe55Leufs*6)	Exon 2	Regulatory	1.04
c.1045T>C	p.Ser349Pro	Exon 10	Catalytic	0.94
c.1042C>G	p.Leu348Val	Exon 10	Catalytic	0.88
c.441+5G>T	p.(?)	IVS 4		0.86
c.331C>T	p.(Arg111*)	Exon 3	Regulatory	0.81
c.814G>T	p.(Gly272*)	Exon 7	Catalytic	0.81
c.781C>T	p.(Arg261*)	Exon 7	Catalytic	0.81
c.168+5G>C	p.(?)	IVS 2		0.69
c.916A>G	p.Ile306Val	Exon 9	Catalytic	0.65
c.1238G>C	p.Arg413Pro	Exon 12	Oligomerization	0.56
c.734T>C	p.Val245Ala	Exon 7	Catalytic	0.55
c.1223G>A	p.Arg408Gln	Exon 12	Catalytic	0.52
c.611A>G	p.(Tyr204Cys)	Exon 6	Catalytic	0.52
c.1055delG	p.(Gly352Valfs*48)	Exon 10	Catalytic	0.50

Calculated on the basis of 14,612 alleles.
Data taken from [26].

are probably representative of the global distribution of metabolic phenotypes among PKU population (TABLE 4).

BH4 responsiveness was reported in 2926 patients (39.4%), 787 (25.0%) of them being classified as responder (blood Phe reduction >30%) or slow responder (blood Phe reduction 20–30%). Most of them (78.8%) belong to the MHP or mild PKU phenotype and 65.8% of non-responders are classic PKU patients (TABLE 4).

The 14,810 alleles affected predominantly the catalytic domain (9885 hits, 67.7%), followed by intronic sites (2351 hits, 16.1%), the regulatory domain (1746 variants, 11.9%) and least the oligomerization domain (628 hits, 4.3%). **The most commonly affected gene locations were exon 12 (4385 hits, 30.0%), exon 7 (3138 hits, 21.5%), IVS 10 (1065 hits, 7.3%) and exon 2 (826 hits, 5.7%).**

Table 4. Summary of the number of different genotypes and responders or non-responders to BH4 challenge in patients tabulated in BIOPKU.

Phenotype	MHP	Mild PKU	Classic PKU	N/A
BH4 responder				
Patients	459	691	248	123
Genotypes	249	298	147	93
Non-responder				
Patients	21	266	949	169
Genotypes	18	173	367	108
Not tested				
Patients	681	825	2373	600
Genotypes	329	394	657	309

MHP: Mild hyperphenylalaninemia; N/A: Not reported; PKU: Phenylketonuria. Data taken from [26].

Genotype–phenotype correlation

One of the major benefits expected from the PKU patients' genotyping is a good genotype–phenotype correlation in order to provide the physician with an effective prognostic interpretive tool. However, since the levels of complexity for correlations with phenotype are manifold [2], the studies on this relationship have been conducted mainly connecting genotype to biochemical phenotype. After *in vitro* expression of several PAH variants [75], many studies have been performed to establish the degree of genotype–phenotype correlation. The biochemical phenotype used in different correlation studies was the highest Phe value before treatment, Phe overloading test or Phe tolerance. The severity of variants was defined by a correlation to patients' phenotype, or in according to predicted residual activity (PRA) provided by *in vitro* experiments or analyzing the variant effects on protein structure.

In 1997, Kayaalp *et al.* [76] compared the patients' phenotype, considering Phe at birth or pretreatment Phe value or Phe tolerance, with genotype when one variant in homozygous or one variant in hemizygous constellation (i.e., in compound heterozygosis with a null variant) was present. The authors observed that 11 variants lied in different phenotypic classes. Further, comparing the variants with the available PRA, the authors observed some inconsistencies too. They explained this as the effect of heteroallelism that could not be evaluated by *in vitro* expression. The variants reported with major inconsistencies were p.Ile65Thr and p.Tyr414Cys.

A further European multicenter study [77] assigned the variant severity considering their highly probable deleterious effect due to their nature (stop variants, frame shift variants or variants affecting the invariant AT-TG consensus splice sequence). For the missense variants, they assigned the activity by PRA values and evaluated the phenotypic effect of other variants when present in hemizygous state and concordantly in more than one patient. Using these criteria, the authors found a 79% of concordance and proposed a

scale to evaluate the severity of variants assigning them an arbitrary value (AV) following this scheme: AV = 1 classical, AV = 2 moderate, AV = 4 mild, AV = 8 non-PKU-HPA. In homozygotes and in compound heterozygotes, the resultant phenotype can be calculated by the sum of the two scores. Inconsistencies were also observed in this work, attributed the majority to lack of standardization among the various participating centers. However, some variants were more frequently reported in discordant genotype–phenotype associations: p.Ile65Thr, p.Arg261Gln, p.Arg158Gln, p.Leu48Ser and p.Tyr414Cys.

Several other studies were performed, although in selected populations using the AV scale and the prevalent observation was that the genotype is a good predictor of the phenotype in all comparative methods used; nevertheless, some inconsistencies were still reported often against the same variants [78,79]. The highest degree of concordance was found for severe or null variants when present in homozygosis or in compound heterozygosis with another null variant, although in this case also few exceptions were reported (c.1066-11G>A, p.Arg408Trp).

After the determination of PAH crystal structure, a different approach to deal with effect of variants was possible. The phenotypic effect of variants was evaluated by their position within the protein functional domains and their possible effect on oligomerization and protein stability [80]. By this approach and by expression studies, some authors [81] found an explanation for some inconsistencies previously reported, ascribing this effect to folding amenable to *in vitro* modulation. The p.Ile65Thr and p.Arg261Gln variants are overrepresented in patients with discordant phenotypes and exhibited a very striking response to temperature in the expression system used.

Characterized missense variants tabulated in the PAHvdb (see above) using FoldX, SIFT, Polyphen-2 and SNPs3D algorithms and calculated the APV [11]. Subsequently, values for residual PAH activity (%PAH) and APV were correlated with patients' phenotype and BH4 response and found a fully quantitative relationship between mutant homotetrameric stability, residual PAH activity and Phe, making %PAH, FoldX and SNP3d as reliable predictors of hemizygous phenotype and supporting misfolding hypothesis in PKU. Although heterotrimer stability was difficult to estimate, compound heterozygous phenotype and BH4 response were predicted from genotype, APV and %PAH.

Primary prevention

The PKU adult patients who are planning a pregnancy often ask about the risk factors associated with having a child with this disease. The same request originates from patient's relatives, regarding healthy carrier for the disease. The *a priori* risk in this case can be calculated considering the fact that in Caucasians a general risk to be carrier is about 1/50. Nevertheless, a better risk assessment could be obtained by performing a genetic test in carrier partners. However, the PAH sequencing analysis conducted in subjects of general population can show some limits: the detection rate is lower than 100% and there is

the chance to not identify the variant; there is the possibility to identify new sequence variants of not easy pathogenic definition. These reasons could be useful for each screening center to create a panel of more frequent variants in the belonging population to be analyzed in a subject of general population that have a couple risk *a priori* higher than in a couple of general population. In this event, the test should be preceded and followed by genetic counsel.

Expert commentary

Advances made over the past 10–20 years in basic science led to the emergence of new treatment options of clinically proven benefit for some PKU patients. BH₄ (sapropterin dihydrochloride) has been approved by the US FDA for the management of PKU and by the EMA for the management of PKU and BH₄-deficiency patients. Treatment with sapropterin has resulted in clinically significant and sustained reductions in blood Phe concentrations and increases in dietary Phe tolerance in well-designed clinical studies in patients with PKU, with or without concurrent treatment with a Phe-restricted diet. Sapropterin is an alternative treatment option for approximately 20–30% of PKU patients, depending on the local frequency of mild or severe genotypes. Most responders to sapropterin have mild-to-moderate PKU phenotypes, but a few patients with classic PKU have also responded to oral administration of BH₄. As a general rule, all patients detected with HPA need to be carefully investigated and their phenotype estimated. Based on the phenotype, individual tolerance for the daily protein intake, and local or national guidelines decision should be made to treat or not to treat. Those patients needing treatment should be further investigated for a potential pharmacological intervention. This includes a BH₄ challenge and genotyping.

Investigation of the effects of variant on enzyme activity by *in vitro* functional assays links the patient data with protein properties. The pathogenic effects of more than 100 variant PAH proteins have been investigated in detailed biochemical studies for a better understanding of relationships between mutant biochemistry and PKU disease presentation (for details see BIOPKU database). All described expression systems lead to the production of homomeric forms of PAH protein. However, most PKU patients are compound heterozygotes carrying different variants and two-hybrid analysis may allow the semi-quantitative assessment of protein–protein interactions in living cells. The interaction of two mutated subunits may exhibit a different response to BH₄ and different molecular behavior as the single homozygous variant alone. Interallelic complementation describes the interaction between two mutant alleles at a given locus to form a multimeric protein. A particular combination of two alleles may produce a more (negative complementation) or less (positive complementation) severe phenotype than expected from the homoallelic counterpart [10]. Thus, complementation may lead to correction of misfolding of one of the monomers. This phenomenon of positive or negative complementation contributes to the phenotypic severity in PKU and may explain the absence of BH₄ response in patients carrying two responsive alleles.

The last scientific and technological advancements and their translation into clinical medicine let us think to be at the dawn of great transformation: the next-generation sequencing on panels of target genes will be routine in molecular laboratory in a few years and this will lead to faster and more accurate phenotypic definition; new therapies are being tested on multiple fronts and in a few years the therapy may be personalized for each patient depending on his/her genotype. In PKU, the treatment can be with BH₄ if variants generate a mild phenotype, or, in subjects with severe stop variants, with new read-through drugs. PKU has always been a paradigmatic example for genetic and metabolic diseases, for its Mendelian inheritance, for preventive medicine through neonatal screening, for its allelic heterogeneity and complexity of the phenotypic expression and will continue to be so in the future.

Five-year view

Next-generation assays: Next-generation sequencing is a new technology that is increasingly used in molecular laboratory. This method allows a deep sequencing at different levels: it can analyze the entire exome, or a more selected panel of target genes, usually associated with a phenotype with genetic heterogeneity, or can analyze an entire gene including its intronic sequence. This technology possesses many advantages: the increased diagnostic sensitivity, the shorter analysis time and the cost, which is much lower than Sanger sequencing while analyzing one gene or of a panel of target genes. However, the major disadvantage is the need of new variant validation [82].

Recently, some *authors* have shown that it is possible to reach 100% of detection rate by analyzing hyperphenylalaninemic samples using a panel of four target genes (*PAH*, *GCH1*, *PTS* and *QDPR*) in a system based on high-throughput multiplex-targeted sequencing. The analysis also includes large genomic rearrangements, indels and deep intronic variants. This strategy allows identifying, with high sensitivity, all variants causing hyperphenylalaninemia without the need for additional specific assays (i.e., multiple ligations probe amplification for deletions/duplications) [83].

Genetic neonatal screening: Genotyping is currently used in newborn screening as confirmation testing. Nevertheless, the potentiality of the next-generation sequencing suggests its application as first screening test for many metabolic diseases. Virtually this technique is ready to be applied in this field: its turnaround time is decreased to 2–3 days and can be reliably performed by using DBSs. However a lot of clinical and ethical problems should be taken into account. It is necessary to consider the possibility of finding variants that define the carrier status for many diseases or variants that can increase the risk for other diseases. For these reasons, nowadays, the American Colleague of Medical Genetics has opposed its use as a first-tier approach to newborn screening [84].

Research on new therapies: PKU is a disorder well treatable by a low Phe diet. However, the long life diet could cause poor compliance, justifying research on therapeutic alternatives. The discovery of BH₄ effects on mutated PAH enzyme activity has contributed to the emergence of a category of ‘misfolding diseases’. This category includes disease in which the variants cause a misfolding in

the protein that lowers its activity. By using chaperones, the mutated protein increases its stability and consequently its activity. Starting from this observation, the search for other compounds with chaperones activity goes ahead following two main strategies: high-throughput screening of diversity libraries and through computational screening [44]. Several compounds have been identified and tried in the PKU mice model. One of them, benzylhydantoin, shows a reduction twice as effective as BH₄ [85].

Phenylalanine ammonia lyase (PAL) is a vegetal enzyme that catalyzes a reaction converting L-Phe to ammonia and trans-cinnamic acid. Its use for lowering Phe blood concentration is considered promising. Studies on animal models show how some formulations (e.g., PEG-PAL) of the enzyme can lower Phe concentration of 40%. Human trials are underway [86] and open-label, multicenter, Phase I dose-escalation trial was found to be fairly safe and well tolerated in 25 adult patients with PKU [87].

Gene therapy, initially thought very promising, actually turned out to be difficult in its clinical applications and there were important issues to be addressed before a gene therapy strategy could be translated into clinical practice, particularly for PKU. On the other hand, in the past years, progresses have been made in variant-targeted treatments. The discovery that aminoglycosides induce stop variant read-through was first reported in CFTR in which protein expression was restored treating cells with low doses of aminoglycoside antibiotics [88]. Stop variants are considered severe as the presence of stop

codon initiate a mechanism that induces degradation of mRNA (nonsense-mediated decay) and the resulting absence of protein. The aminoglycoside acts inducing the binding of a similar tRNA allowing translation to proceed maintaining the correct reading frame. The research has been focused on compounds that act similarly to aminoglycosides but with fewer side effects. A compound called PTC124 (Ataluren) was tried; it has a strong read-through capacity with no toxic effect [89]. The effects of read-through drugs have also been studied in metabolic disorders [90]. In PKU, *in vitro* experiment demonstrated in two mammalian cell lines (COS-7 and HEK293) the restoration of transcription using gentamicin and Geneticin. The recovered enzyme activity is much less than wild type, getting close to a missense variant enzyme activity and the drug dosage too high for clinical usage. However, results encourage studies on new read-through drugs [91]. The overview of research on new therapies looks very promising and toward personalized medicine. We expect significant results in the near future.

Financial & competing interests disclosure

This work is part of the RD-CONNECT initiative and was supported by the FP7-HEALTH-2012-INNOVATION-1 EU Grant No. 305444 (to NB). The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Key issues

- Advances made over the past 10–20 years in basic science led to new treatments of clinically proven benefit for phenylketonuria patients.
- The dried blood spots showed highly specific and accurate results for both biochemical and molecular testing of patients with hyperphenylalaninemia.
- A large number of variants (>850), many being private, urged establishment of locus-specific and genotype databases and development of new techniques for the molecular diagnosis of phenylketonuria.
- Increasing number of online available tools for variant-severity and protein damage analysis enabled investigation of the genotype–phenotype correlation by the use of prediction algorithms.
- Next-generation sequencing has emerged as a new technology in molecular laboratory. This technology possesses many advantages: the increased diagnostic sensitivity, the shorter analysis time and lower costs, particularly for a panel of target genes.

References

1. Blau N, Van Spronsen FJ, Levy HL. Phenylketonuria. *Lancet* 2010;376:1417-27
2. Scriver CR, Waters PJ. Monogenic traits are not simple: lessons from phenylketonuria. *Trends Genet* 1999;15(7):267-72
3. Fölling Ä. Über Ausscheidung von Phenylbrenztraubensäure in den Harn als Stoffwechsellanomale in Verbindung mit Inbicillität. *Hoppe-Seylers Ztschr Physiol Chem* 1934;227:169
4. Penrose LS. Inheritance of phenylpyruvic amentia (phenylketonuria). *Lancet* 1935;2:192-4
5. Jervis GA. Phenylpyruvic oligophrenia deficiency of phenylalanine-oxidizing system. *Proc Soc Exp Biol Med* 1953;82(3):514-15
6. Bickel H, Gerrard JW, Hickmans EM. Influence of phenylalanine intake on phenylketonuria. *Lancet* 1953;2:812-19
7. Guthrie R, Susi A. A simple phenylalanine method for detecting phenylketonuria in large populations of newborn infants. *Pediatrics* 1963;32:338-43
8. Lidsky AS, Robson KJ, Thirumalachary C, et al. The PKU locus in man is on chromosome 12. *Am J Hum Genet* 1984;36(3):527-33
9. Scriver CR. The PAH gene, phenylketonuria, and a paradigm shift. *Hum Mutat* 2007;28:831-45
10. Heintz C, Cotton RG, Blau N. Tetrahydrobiopterin, its mode of action on phenylalanine hydroxylase and importance of genotypes for pharmacological therapy of phenylketonuria. *Hum Mutat* 2013;34(7):927-36

11. Wettstein S, Underhaug J, Perez B, et al. Linking genotypes database with locus-specific database and genotype-phenotype correlation in phenylketonuria. *Eur J Hum Genet* 2014. doi:10.1038/ejhg.2014.114
12. Beasley MG, Costello PM, Smith I. Outcome of treatment in young adults with phenylketonuria detected by routine neonatal screening between 1964 and 1971. *Q J Med* 1994;87(3):155-60
13. Waisbren SE, Noel K, Fahrback K, et al. Phenylalanine blood levels and clinical outcomes in phenylketonuria: a systematic literature review and meta-analysis. *Mol Genet Metab* 2007;92(1-2):63-70
14. Walter JH, White FJ, Hall SK, et al. How practical are recommendations for dietary control in phenylketonuria? *Lancet* 2002;360:55-7
15. Antshel KM, Waisbren SE. Developmental timing of exposure to elevated levels of phenylalanine is associated with ADHD symptom expression. *J Abnorm Child Psychol* 2003;31(6):565-74
16. Camp KM, Parisi MA, Acosta PB, et al. Phenylketonuria Scientific Review Conference: state of the science and future research needs. *Mol Genet Metab* 2014. [Epub ahead of print]
17. Surtees R, Blau N. Neurochemistry of phenylketonuria. *Eur J Pediatr* 2000;159:109-13
18. van Spronsen FJ, Hoeksma M, Reijngoud DJ. Brain dysfunction in phenylketonuria: is phenylalanine toxicity the only possible cause? *J Inherit Metab Dis* 2009;32(1):46-51
19. Feillet F, van Spronsen FJ, MacDonald A, et al. Challenges and pitfalls in the management of phenylketonuria. *Pediatrics* 2010;126:333-41
20. van Spronsen FJ, Burgard P. The truth of treating patients with phenylketonuria after childhood: the need for a new guideline. *J Inherit Metab Dis* 2008;31(6):673-9
21. MacDonald A. Diet and compliance in phenylketonuria. *Eur J Pediatr* 2000;159(Suppl 2):S136-41
22. National institutes of health consensus development conference statement: phenylketonuria: screening and management, October 16-18, 2000. *Pediatrics* 2001;108(4):972-82
23. Albrecht J, Garbade SF, Burgard P. Neuropsychological speed tests and blood phenylalanine levels in patients with phenylketonuria: a meta-analysis. *Neurosci Biobehav Rev* 2009;33(3):414-21
24. Güttler F. Hyperphenylalaninemia: diagnosis and classification of the various types of phenylalanine hydroxylase deficiency in childhood. *Acta Paediatr Scand Suppl* 1980;280:1-80
25. Blau N, Hennermann JB, Langenbeck U, Lichter-Konecki U. Diagnosis, classification and genetics of phenylketonuria and tetrahydrobiopterin (BH4) deficiencies. *Mol Genet Metab* 2011;104:S2-9
26. Blau N, Yue WW, Perez B. International database of patients and genotypes causing HPA/PKU incl. BH4-responsive phenotype. In: Available at www.biopku.org/home/biopku.asp [Last accessed 17 April 2014]
27. Weglage J, Wiedermann D, Denecke J, et al. Individual blood-brain barrier phenylalanine transport determines clinical outcome in phenylketonuria. *Ann Neurol* 2001;50(4):463-7
28. Waters PJ. How PAH gene mutations cause hyper-phenylalaninemia and why mechanism matters: insights from in vitro expression. *Hum Mutat* 2013;21(4):357-69
29. Kure S, Hou DC, Ohura T, et al. Tetrahydrobiopterin-responsive phenylalanine hydroxylase deficiency. *J Pediatr* 1999;135:375-8
30. Bernegger C, Blau N. High frequency of tetrahydrobiopterin-responsiveness among hyperphenylalaninemias: a study of 1919 patients observed from 1988 to 2002. *Mol Genet Metab* 2002;77:304-13
31. Fiege B, Blau N. Assessment of tetrahydrobiopterin (BH4)-responsiveness in phenylketonuria. *J Pediatr* 2007;150:627-30
32. Muntau AC, Röschinger W, Habich M, et al. Tetrahydrobiopterin as an alternative treatment for mild phenylketonuria. *N Engl J Med* 2002;347:2122-32
33. Blau N, Bélanger-Quintana A, Demirkol M, et al. Optimizing the use of sapropterin (BH4) in the management of phenylketonuria. *Mol Genet Metab* 2009;96:158-63
34. Levy H, Milanowski A, Chakrapani A, et al. Efficacy of sapropterin dihydrochloride (tetrahydrobiopterin, 6R-BH4) for reduction of phenylalanine concentration in patients with phenylketonuria: a phase III randomized placebo-controlled study. *Lancet* 1994;370:504-10
35. Keil S, Anjema K, van Spronsen J, et al. Long-term follow-up and outcome of phenylketonuria patients on sapropterin: a retrospective study. *Pediatrics* 2013;131(6):e1881-8
36. Fiege B, Bonafé L, Ballhausen D, et al. Extended tetrahydrobiopterin loading test in the diagnosis of cofactor-responsive phenylketonuria: a pilot study. *Mol Genet Metab* 2005;86(Suppl 1):91-95
37. Anjema K, van Rijn M, Hofstede FC, et al. Tetrahydrobiopterin responsiveness in phenylketonuria: prediction with the 48-hour loading test and genotype. *Orphanet J Rare Dis* 2013;8(1):103
38. Shintaku H, Kure S, Ohura T, et al. Long-term treatment and diagnosis of tetrahydrobiopterin-responsive hyperphenylalaninemia with a mutant phenylalanine hydroxylase gene. *Pediatr Res* 2004;55:425-30
39. Karacic I, Meili D, Sarnavka V, et al. Genotype-predicted tetrahydrobiopterin (BH4)-responsiveness and molecular genetics in Croatian patients with phenylalanine hydroxylase (PAH) deficiency. *Mol Genet Metab* 2009;97(3):165-71
40. Dobrowolski SF, Heintz C, Miller T, et al. Molecular genetics and impact of residual in vitro phenylalanine hydroxylase activity on tetrahydrobiopterin-responsiveness in Turkish PKU population. *Mol Genet Metab* 2011;10:116-21
41. Zurflüh MR, Zschocke J, Lindner M, et al. Molecular genetics of tetrahydrobiopterin responsive phenylalanine hydroxylase deficiency. *Hum Mutat* 2008;29(1):167-75
42. Trefz FK, Scheible D, Götz H, Frauendienst-Egger G. Significance of genotype in tetrahydrobiopterin-responsive phenylketonuria. *J Inherit Metab Dis* 2009;32:22-6
43. Quirk ME, Dobrowolski SF, Nelson BE, et al. Utility of phenylalanine hydroxylase genotype for tetrahydrobiopterin responsiveness classification in patients with phenylketonuria. *Mol Genet Metab* 2012;107(1-2):31-6
44. Underhaug J, Aubi O, Martinez A. Phenylalanine hydroxylase misfolding and pharmacological chaperones. *Curr Top Med Chem* 2013;12(22):2534-45
45. Werner ER, Blau N, Thöny B. Tetrahydrobiopterin: biochemistry and pathophysiology. *Biochem J* 2011;438:397-414
46. Heintz C, Troxler H, Martinez A, et al. Quantification of phenylalanine hydroxylase activity by isotope-dilution liquid chromatography-electrospray ionization tandem mass spectrometry. *Mol Genet Metab* 2012;105(4):559-65
47. Martinez A, Knappskog PM, Olafsdottir S, et al. Expression of recombinant human phenylalanine hydroxylase as fusion protein in *Escherichia coli* circumvents proteolytic

- degradation by host cell proteases - Isolation and characterization of the wild-type enzyme. *Biochem J* 1995;306(Part 2): 589-97
48. Shiman R, Gray DW, Hill MA. Regulation of rat liver phenylalanine hydroxylase. I. Kinetic properties of the enzyme's iron and enzyme reduction site. *J Biol Chem* 1994; 269(40):24637-46
 49. Shiman R, Xia T, Hill MA, Gray DW. Regulation of rat liver phenylalanine hydroxylase. II. Substrate binding and the role of activation in the control of enzymatic activity. *J Biol Chem* 1994; 269(40):24647-56
 50. Kobe B, Jennings IG, House CM, et al. Structural basis of autoregulation of phenylalanine hydroxylase. *Nat Struct Biol* 1999;6(5):442-8
 51. DiLella AG, Marvit J, Lidsky AS, et al. Tight linkage between a splicing mutation and a specific DNA haplotype in phenylketonuria. *Nature* 1986;322(6082): 799-803
 52. The Nucleotide Database. 2014. Available from: www.ncbi.nlm.nih.gov/nucleotide/ [Last accessed 17 April 2014]
 53. The Consensus CDS (CCDS) Database. 2014. Available at www.ncbi.nlm.nih.gov/projects/CCDS/ [Last accessed 17 April 2014]
 54. Konecki DS, Wang Y, Trefz FK, et al. Structural characterization of the 5' regions of the human phenylalanine hydroxylase gene. *Biochemistry* 1992;31(35):8363-8
 55. Gene Database. 2014. Available at www.ncbi.nlm.nih.gov/gene/ [Last accessed 17 April 2014]
 56. Blau N, Yue WW, Perez B. Phenylalanine hydroxylase gene locus-specific database. In: Available at www.biopku.org/home/pah.asp [Last accessed 17 April 2014]
 57. den Dunnen JT, Antonarakis SE. Nomenclature for the description of sequence variants. 2014. Available at www.hgvs.org/mutnomen/ [Last accessed 17 April 2014]
 58. Online mendelian inheritance in man (OMIM). 2014. Available at www.omim.org/ [Last accessed 17 April 2014]
 59. Cooper DN, et al. The human gene mutation database. 2014. Available at www.hgmd.org/ [Last accessed 17 April 2014]
 60. Horatis R. Disease centered central mutation databases. 2014. Available at www.hgvs.org/dblist/disease.html [Last accessed 17 April 2014]
 61. Robson KJ, Beattie W, James RJ, et al. Sequence comparison of rat liver phenylalanine hydroxylase and its cDNA clones. *Biochemistry* 1984;23(24): 5671-5
 62. Lidsky AS, Ledley FD, DiLella AG, et al. Extensive restriction site polymorphism at the human phenylalanine hydroxylase locus and application in prenatal diagnosis of phenylketonuria. *Am J Hum Genet* 1985; 37(4):619-34
 63. DiLella AG, Marvit J, Brayton K, Woo SL. An amino-acid substitution involved in phenylketonuria is in linkage disequilibrium with DNA haplotype 2. *Nature* 1987; 327(6120):333-6
 64. DiLella AG, Huang WM, Woo SL. Screening for phenylketonuria mutations by DNA amplification with the polymerase chain reaction. *Lancet* 1988;1(8584):497-9
 65. Dworniczak B, Kalaydjieva L, Pankoke S, et al. Analysis of exon 7 of the human phenylalanine hydroxylase gene: a mutation hot spot? *Hum Mutat* 1992;1(2):138-46
 66. Michiels L, François B, Raus J, Vandevyver C. Rapid identification of PKU-associated mutations by multiplex DGGE analysis of the PAH gene. *J Inherit Metab Dis* 1996;19(6):735-8
 67. Brautigam S, Kujat A, Kirst P, et al. DHPLC mutation analysis of phenylketonuria. *Mol Genet Metab* 2003; 78(3):205-10
 68. Kozak L, Hrabincova E, Kintz J, et al. Identification and characterization of large deletions in the phenylalanine hydroxylase (PAH) gene by MLPA: evidence for both homologous and non-homologous mechanisms of rearrangement. *Mol Genet Metab* 2006;89(4):300-9
 69. Cali F, Ruggeri G, Vinci M, et al. Exon deletions of the phenylalanine hydroxylase gene in Italian hyperphenylalaninemics. *Exp Mol Med* 2010;42(2):81-6
 70. Gamez A, Pérez B, Ugarte M, Desviat LR. Expression analysis of phenylketonuria mutations. Effect on folding and stability of the phenylalanine hydroxylase protein. *J Biol Chem* 2000;275(38):29737-42
 71. Pey AL, Pérez B, Desviat LR, et al. Mechanisms underlying responsiveness to tetrahydrobiopterin in mild phenylketonuria mutations. *Hum Mutat* 2004;24(5):388-99
 72. Sarkar G, Sommer SS. Access to a messenger RNA sequence or its protein product is not limited by tissue or species specificity. *Science* 1989;244(4902):331-4
 73. Ramus SJ, Forrest SM, Cotton RG. Illegitimate transcription of phenylalanine hydroxylase for detection of mutations in patients with phenylketonuria. *Hum Mutat* 1992;1(2):154-8
 74. Heintz C, Dobrowolski SF, Andersen HS, et al. Splicing of phenylalanine hydroxylase (PAH) exon 11 is vulnerable: molecular pathology of mutations in PAH exon 11. *Mol Genet Metab* 2012;106:403-11
 75. Waters PJ, Parniak MA, Nowacki P, Scriver CR. In vitro expression analysis of mutations in phenylalanine hydroxylase: linking genotype to phenotype and structure to function. *Hum Mutat* 1998;11(1):4-17
 76. Kayaalp E, Treacy E, Waters PJ, et al. Human phenylalanine hydroxylase mutations and hyperphenylalaninemia phenotypes: a metanalysis of genotype-phenotype correlations. *Am J Hum Genet* 1997;61(6):1309-17
 77. Guldberg P, Rey F, Zschocke J, et al. A European multicenter study of phenylalanine hydroxylase deficiency: classification of 105 mutations and a general system for genotype-based prediction of metabolic phenotype. *Am J Hum Genet* 1998;63(1):71-9
 78. Zhu T, Ye J, Han L, et al. Variations in genotype-phenotype correlations in Phenylalanine hydroxylase deficiency in Chinese Han Population. *Gene* 2013; 529(1):80-7
 79. Polak E, Ficek A, Radvanszky J, et al. Phenylalanine hydroxylase deficiency in the Slovak population: genotype-phenotype correlations and genotype-based predictions of BH4-responsiveness. *Gene* 2013;526(2): 347-55
 80. Jennings IG, Cotton RG, Kobe B. Structural interpretation of mutations in phenylalanine hydroxylase protein aids in identifying genotype-phenotype correlations in phenylketonuria. *Eur J Hum Genet* 2000;8(9):683-96
 81. Pey AL, Desviat LR, Gámez A, et al. Phenylketonuria: genotype-phenotype correlations based on expression analysis of structural and functional mutations in PAH. *Hum Mutat* 2003;21(4):370-8
 82. Metzker ML. Sequencing technologies - the next generation. *Nat Rev Genet* 2010;11(1): 31-46
 83. Trujillano D, Perez B, González J, et al. Accurate molecular diagnosis of phenylketonuria and tetrahydrobiopterin-deficient hyperphenylalaninemics using high-throughput targeted sequencing. *Eur J Hum Genet* 2014;22(4):528-34

84. Landau YE, Lichter-Konecki U, Levy HL. Genomics in newborn screening. *J Pediatr* 2014;164(1):14-19
85. Santos-Sierra S, Kirchmair J, Perna AM, et al. Novel pharmacological chaperones that correct phenylketonuria in mice. *Hum Mol Genet* 2012;21(8):1877-87
86. Belanger-Quintana A, Burlina A, Harding CO, Muntau AC. Up to date knowledge on different treatment strategies for phenylketonuria. *Mol Genet Metab* 2011(104 Suppl):S19-25
87. Longo N, Harding CO, Burton BK, et al. Single-dose, subcutaneous recombinant phenylalanine ammonia lyase conjugated with polyethylene glycol in adult patients with phenylketonuria: an open-label, multicentre, phase 1 dose-escalation trial. *Lancet* 2014. doi:10.1016/S0140-6736(13)61841-3
88. Howard M, Frizzell RA, Bedwell DM. Aminoglycoside antibiotics restore CFTR function by overcoming premature stop mutations. *Nat Med* 1996;2(4):467-9
89. Welch EM, Barton ER, Zhuo J, et al. PTC124 targets genetic disorders caused by nonsense mutations. *Nature* 2007; 447(7140):87-91
90. Perez B, Rodríguez-Pombo P, Ugarte M, Desviat LR. Readthrough strategies for therapeutic suppression of nonsense mutations in inherited metabolic disease. *Mol Syndromol* 2012;3(5):230-6
91. Ho G, Reichardt J, Christodoulou J. In vitro read-through of phenylalanine hydroxylase (PAH) nonsense mutations using aminoglycosides: a potential therapy for phenylketonuria. *J Inher Metab Dis* 2013;36(6):955-9