Review article

The phenylketonuria locus: current knowledge about alleles and mutations of the phenylalanine hydroxylase gene in various populations

David S. Konecki and Uta Lichter-Konecki

Universitäts-Kinderklinik, Im Neuenheimer Feld 150, W-6900 Heidelberg, Federal Republic of Germany

Received December 17, 1990 / Revised February 10, 1991

Summary. The hyperphenylalaninemic disorders of classic phenylketonuria (PKU), mild phenylketonuria, and hyperphenylalaninemia (HPA), result from a deficiency of the hepatic enzyme phenylalanine hydroxylase (PAH) or its cofactor (tetrahydrobiopterin). Use of the complementary DNA of this enzyme has allowed the establishment of a restriction fragment length polymorphism (RFLP) haplotype-analysis system. This haplotype analysis system provides the means for determination of mutant PAH alleles in most affected families and is the basis for mutational analysis of the PKU locus. This review is focused on two major areas of current PKU research: (1) the use of DNA haplotype analysis in the study of the population genetics of PAH deficiency, and (2) the study of genotypes, and their various combinations, as a means of explaining and predicting the phenotypic variability observed for the disorders of PAH deficiency.

Introduction

Classic phenylketonuria (PKU) is an autosomal recessive human genetic disorder caused by a deficiency of the hepatic enzyme phenylalanine hydroxylase (phenylalanine 4-monooxygenase, E.C. 1.14.16.1), which catalyzes the hydroxylation of phenylalanine to tyrosine by molecular oxygen in the presence of the cofactor tetrahydrobiopterin (Kaufman 1976). This inborn error of amino acid metabolism causes postnatal brain damage and severe mental retardation in untreated affected children. A deficiency in this enzyme results in the excretion of large quantities of phenylpyruvate in the urine, an accumulation of phenylalanine in the blood resulting in hyperphenylalaninemia, and the abnormal formation of a myelin sheath around neuronal axons in the central nervous system (Kaufman 1976). While classic PKU is the best known of the hyperphenylalaninemic disorders resulting from PAH deficiency, less severe forms of the disease, mild PKU and hyperphenylalaninemia (HPA), also exist. Clinically and biochemically, these disorders are classified according to pretreatment plasma phenylalanine levels, phenylalanine tolerance (Göttler 1984), residual PAH enzyme activity (determined in liver biopsy material or calculated from phenylalanine tolerance data), and response to protein load (Trefz et al. 1985).

PKU is the most common clinically important inborn error of amino acid metabolism, with an average incidence of about 1 in 10,000 Caucasian neonates (Bickel et al. 1981). This frequency varies throughout Europe, ranging from approximately 1:2600 in Turkey (Özalp et al. 1986) to about 1:3000 in Sweden (Veale 1980). In Oriental populations the incidence of PKU ranges from about 1:16000 in China (Daiger et al. 1989b) to approximately 1:119,000 in Japan (Akio and Wada 1988). These incidence values reflect heterozygote frequencies ranging from about 1:34 in Turkey (this value takes into account the high level of consanguinity reported by Özalp et al. 1986) to approximately 1:173 in Japan.

The disease state of PKU was first described by Fölling in 1934 and identified as an autosomal recessive genetic disorder by Penrose in 1955. Since that time PKU has been the focus of extensive research efforts. Major contributions by Jervis (1947 and 1953), Kaufman (1957, 1958, and 1976) and other biochemists have resulted in the delineation of the complex reaction resulting in the hydroxylation of phenylalanine to tyrosine. In 1954, Bickel et al. introduced a dietary treatment strategy for PKU, which was the first demonstration that hereditary disorders of amino acid metabolism could be corrected by dietary management. The success of this treatment, which must be implemented soon after birth to be effective, provided the impetus to develop a screening method to identify PKU in newborn patients. The resulting bacterial inhibition assay, reported by Guthrie and Susi

Offprint requests to: D.S. Konecki
in 1963, is specific, inexpensive, convenient for analyzing large numbers of samples, and has served as a newborn screening prototype for other metabolic diseases. The newborn screening programs for phenylalanine hydroxylase deficiency have revealed a large variety of clinical and biochemical phenotypes. Only now are we beginning to understand the molecular basis of this phenomenon.

The prerequisite for the molecular genetic analysis of the heterogeneity associated with this disease was the isolation of the complementary DNA coding for the PAH enzyme by Prof. S. Woo's group in Houston (Woo et al. 1983). Through the use of in situ hybridization, with a full-length PAH-cDNA as a probe (Kwok et al. 1985), the position of the PAH gene and the PKU locus in man was mapped to the long arm of chromosome 12, position 12q22–24.1 (Lidsky et al. 1985a). Eukaryotic, as well as prokaryotic, expression studies with this cDNA have conclusively demonstrated that a single mRNA species, about 2500 nucleotides in length, contains the genetic information necessary to code for a functional PAH enzyme (Ledley et al. 1985).

Subsequently, the results of studies delineating the organization of the PAH gene were published by DiLella et al. (1986a). Based on the observation that specific mutations are in strong linkage disequilibrium with certain haplotypes in β-thalassemia (Antonarkis et al. 1982; Kazazian et al. 1984), Southern analyses were initiated, using the radiolabelled PAH-cDNA as a probe, to determine the polymorphic nature of the PAH gene. These studies resulted in the establishment of a restriction fragment length polymorphism (RFLP) haplotype-analysis system (Lidsky et al. 1985b; Woo 1988). This system demonstrated conclusively the segregation of mutant PAH genes in the majority of PKU families and is the basis of mutation analysis at this locus. Consequently, this RFLP haplotype-analysis system has been applied to many different populations. Analysis of different mutant alleles has resulted in the detection of some 31 different DNA alterations.

**DNA haplotype analysis and allelic distributions**

The PAH-cDNA detects ten RFLPs (Lidsky et al. 1985b) at the chromosomal PAH locus, eight of which have been employed for DNA haplotype determination. Currently, more than 46 haplotypes have been observed in the populations studied (Woo 1988; Chen et al. 1989; Daiger et al. 1989a, b; Trefz et al. 1990) out of the 1152 haplotypes theoretically possible for the eight utilized RFLPs, (one three-allelic and seven two-allelic).

In the first population investigated, the Danish population, an association between normal and mutant genes and certain haplotypes was observed. About 90% of the PKU chromosomes were associated with four haplotypes, 1 through 4, while haplotypes 2 and 3 were observed predominantly among mutant PAH genes and were rare among normal PAH genes. These two haplotypes, 2 and 3, represented 50% of the mutant PAH genes in Danish, Swedish, and Danish American PKU associated with haplotypes 1 and 4. This led to the hypothesis that there exists an association between certain alleles and specific mutations at the human PAH locus, as had been observed at the β-globin locus with regard to β-thalassemias (Orkin and Kazazian 1984).

Several reports concerning the haplotype distribution at the human phenylalanine hydroxylase locus for various populations have been published (Chakraborty et al. 1987; Aulehla-Scholz et al. 1988; Herrmann et al. 1988; Lichter-Konecki et al. 1988b; Rey et al. 1988; Riess et al. 1988; Chen et al. 1989; Daiger et al. 1989a, b; Hertzberg et al. 1989; Lichter-Konecki et al. 1989a; Stuhmann et al. 1989; Sullivan et al. 1989; Apold et al. 1990; Berthelon et al. 1991; Dianzani et al. 1990a, b; Jaruzel ska et al. 1991; Svensson et al. 1991; Trefz et al. 1990; Zyguj ska et al. 1991; U. Lichter-Konecki, unpublished results). These haplotype distributions show certain patterns throughout the different populations. This distribution of the major mutant and normal haplotypes among various populations (and races), as reported in these references, is presented in Fig. 1 and 2, respectively. Essentially all population-based haplotype analyses at this locus (including Caucasian, Euroasian, Asian, and Polynesian) have detected the presence of haplotypes 1 and 4 among normal (Fig. 2A, B) and mutant PAH alleles (Fig. 1A, B). Populations not following this trend are the Japanese (no normal haplotype 1 alleles detected), Czechoslovakian and Chinese (no mutant haplotype 1 alleles detected in either population). However, relatively few PAH chromosomes have been analyzed for individuals of these three countries. Haplotypes 5 and 7 were the next most frequent normal alleles, while the most frequent mutant alleles were represented by haplotypes 2, 3, 6, and 7. Therefore, haplotypes 1, 4, and 7 have been suggested to be ancient PAH haplotypes predating the divergence of races (Hertzberg et al. 1989), while the other alleles probably emerged from these haplotypes by way of different mutational or recombinational events. The recent identification of two silent (same-sense) mutations (A→G, Gln^{252}→Gln and G→A, Val^{245}→Val; Lichter-Konecki et al. 1990) within the coding region of the PAH gene, and subsequent studies to determine haplotype association, demonstrate conclusively that haplotype 7 is derived from haplotype 4. Our studies have found the Gln^{252} silent mutation in PAH exon 6 to be associated with normal and mutant haplotype 3, 4, and 7 alleles, whereas the Val^{245} same-sense mutation in PAH exon 7 is specifically associated with haplotype 4 alleles. Thus, only haplotypes 1 and 4 can be considered ancient PAH haplotypes. To further test this hypothesis we performed PAH-DNA haplotype analysis of DNA isolated from a chimpanzee and an orang-utan. The chimpanzee shared most, but not all, polymorphic restriction fragments with man, while exhibiting a different pattern of constant bands overall. These results prevented a direct comparison of chimpanzee alleles with human alleles at the PAH locus. The orang-utan possessed a restriction fragment pattern different from humans at its PAH gene locus, indicating variability in the positions of the polymorphic restriction sites between both these species of apes and man (U. Lichter-Konecki, unpublished data).
As shown in Fig. 1A, a decrease in the frequency of the mutant haplotype 2 allele from east to west has been reported in Europe, the frequencies for this allele being highest in Czechoslovakia (68%) and Hungary (55%), then decreasing from East (54%) to West Germany (26%) and into France (17.6%). [The data of Rey et al. (1988) may not reflect the actual distribution in France, since a significant number of North African patients were included in their study.] Also demonstrated in Fig. 1A is a decreasing frequency of mutant haplotype 2 and 3 alleles toward the south of Europe, with an increasing “dominance” of other haplotypes among the mutant alleles, i.e., the haplotype 6 allele in Italy and Turkey (Woo et al. 1991). RFLP haplotype analysis of the Turkish population showed that haplotypes 1 and 4 account for the majority of mutant alleles (52%), whereas haplotypes 2 and 3 together represented only about 3% of the PAH mutant alleles in Turkey. About 40% of the chromosomes carrying mutant PAH genes segregate with PAH haplotype 6 alleles, which were very rare among the normal Turkish chromosomes (Lichter-Konecki et al. 1989a). This haplotype shows a low frequency among normal or mutant chromosomes in northern and western European, as well as in Asian populations. Since the mutant haplotype 6 allele is relatively frequent in Italy, and shows an increasing frequency in Turkey, it may in the future be found to be as dominant a mutant allele in Mediterranean countries, as is the haplotype 3 mutant allele in northern and western Europe.

Due to the apparent low frequency in the populations of China and Japan (approximately 15- to 20-fold less than in most Caucasian populations, Thalhammer 1975),
PKU was once considered to be a disorder of Caucasians. Recent investigations have altered this picture, revealing the incidence of PKU in China (Daiger et al. 1989b) and Japan (Aoki and Wada 1988) to be 1:16000 and 1:119000, respectively. Haplotype analyses among Asians have shown haplotype 4 to account for 74–80% of the mutant alleles (Fig. 1B). Thus far, no mutant haplotype 2 or 3 alleles have been detected in Japan and only one of each in China (Daiger et al. 1989b). These data indicate that haplotypes 2 and 3, and the mutations associated with them, were not present at all until relatively recently in northern and central European populations. The fifth most frequent mutant allele in the German population, the haplotype 7 PKU allele, was also examined in Denmark, France, Norway, Poland, and Sweden, as well as China and Japan. These haplotype distributions, when viewed with the results of the mutations analyzed, indicate certain patterns concerning the distribution of mutations and the relation between PAH haplotypes and specific mutations in different populations.

**Association of mutations and alleles at the PAH gene locus**

The original observation at the β-globin locus that specific mutations were in strong linkage disequilibrium with certain β-thalassemia haplotypes (Orkin and Kazazian 1984) provided the foundation for the approach taken to analyze mutations in the PAH gene resulting in hyperphenylalaninemia. Having observed that haplotype 3 represented 38% of the mutant alleles and only 3% of the normal alleles in the Danish population (Gütter-
<table>
<thead>
<tr>
<th>Type of mutation</th>
<th>Genotype</th>
<th>Effect of mutation</th>
<th>Associated with RFLP haplotype (in population)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Donor splice site</td>
<td>G→A intron 12</td>
<td>Truncated protein $a$</td>
<td>3 (European)</td>
<td>DiLella et al. (1986b)</td>
</tr>
<tr>
<td>(2) Missense</td>
<td>C→T$^b$ exon 12</td>
<td>Arg$^{206}$→Trp$^a$</td>
<td>2 (European)</td>
<td>DiLella et al. (1987)</td>
</tr>
<tr>
<td>(3) Missense</td>
<td>T→C exon 9</td>
<td>Leu$^{311}$→Pro$^a$</td>
<td>1 (German)</td>
<td>John et al. (1990)</td>
</tr>
<tr>
<td>(4) Missense</td>
<td>T→C exon 9</td>
<td>Leu$^{311}$→Pro$^a$</td>
<td>10 (German)</td>
<td>Lichter-Konecki et al. (1988a)</td>
</tr>
<tr>
<td>(5) Missense</td>
<td>T→C exon 9</td>
<td>Leu$^{311}$→Pro$^a$</td>
<td>7 (Greek)</td>
<td>Hofman et al. (1989)</td>
</tr>
<tr>
<td>(6) Missense</td>
<td>G→A$^b$ exon 7</td>
<td>Glu$^{280}$→Lys$^a$</td>
<td>38 (N. African, one French)</td>
<td>Lyonnet et al. (1989)</td>
</tr>
<tr>
<td>(7) Missense</td>
<td>G→A$^b$ exon 7</td>
<td>Glu$^{280}$→Lys$^a$</td>
<td>4 (Algerians)</td>
<td>Abadie et al. (1989)</td>
</tr>
<tr>
<td>(8) Missense</td>
<td>G→A$^b$ exon 7</td>
<td>Glu$^{280}$→Lys$^a$</td>
<td>1 (Caucasian)</td>
<td>Okano et al. (1990b)</td>
</tr>
<tr>
<td>(9) Missense</td>
<td>C→T$^b$ exon 3</td>
<td>Arg$^{111}$→Ter$^a$</td>
<td>4 (Chinese)</td>
<td>Wang et al. (1989a, 1990b)</td>
</tr>
<tr>
<td>(10) Missense</td>
<td>A→G exon 1</td>
<td>Met$^{1}$→Val$^a$</td>
<td>2 (French Canadian)</td>
<td>John et al. (1989)</td>
</tr>
<tr>
<td>(11) Missense</td>
<td>G→A$^b$ exon 7</td>
<td>Arg$^{281}$→Gln$^a$</td>
<td>1 (European)</td>
<td>Okano et al. (1990a, c)</td>
</tr>
<tr>
<td>(12) Missense</td>
<td>G→A$^b$ exon 7</td>
<td>Arg$^{281}$→Gln$^a$</td>
<td>4 (European)</td>
<td>Okano et al. (1990a, c)</td>
</tr>
<tr>
<td>(13) Deletion</td>
<td>exon 11</td>
<td>Leu$^{264}$ deleted</td>
<td>5 (Swedish)</td>
<td>Dworniczak et al. (1989)</td>
</tr>
<tr>
<td>(14) Missense</td>
<td>T→G exon 8</td>
<td>Phe$^{260}$→Cys$^a$</td>
<td>? (Caucasian)</td>
<td>Okano et al. 1989, 1990c</td>
</tr>
<tr>
<td>(15) Missense</td>
<td>C→T$^b$ exon 7</td>
<td>Arg$^{212}$→Trp$^a$</td>
<td>1 (Caucasian)</td>
<td>Abadie et al. (1989)</td>
</tr>
<tr>
<td>(16) Missense</td>
<td>C→T$^b$ exon 7</td>
<td>Pro$^{261}$→Leu$^a$</td>
<td>1 (+4) (Caucasian)</td>
<td>Okano et al. 1989, 1990c</td>
</tr>
<tr>
<td>(17) Missense</td>
<td>A→G exon 6</td>
<td>Tyr$^{269}$→Cys$^a$</td>
<td>4 (Chinese)</td>
<td>Wang et al. (1989b)</td>
</tr>
<tr>
<td>(18) Missense</td>
<td>G→C$^b$ exon 12</td>
<td>Arg$^{131}$→Pro$^a$</td>
<td>4 (Chinese)</td>
<td>Wang et al. (1989b)</td>
</tr>
<tr>
<td>(19) Missense</td>
<td>A→G exon 12</td>
<td>Tyr$^{414}$→Cys$^a$</td>
<td>4 (Caucasian)</td>
<td>Okano et al. (1990c)</td>
</tr>
<tr>
<td>(20) Deletion</td>
<td>5' end of gene</td>
<td>NA$^c$ (Scots)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(21) Deletion</td>
<td>exon 3</td>
<td>NA$^c$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(22) Deletion</td>
<td>3' end of gene</td>
<td>NA$^c$ (Japanese)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(23) Mis-sense</td>
<td>T→C exon 7</td>
<td>Leu$^{255}$→Ser</td>
<td>?</td>
<td>Okano et al. 1989, 1990c</td>
</tr>
<tr>
<td>(26) Silent</td>
<td>A→T exon 11</td>
<td>Val$^{399}$→Val</td>
<td>4 (Chinese)</td>
<td>Huang et al. (1991)</td>
</tr>
<tr>
<td>(27) Acceptor splice site</td>
<td>G→A intron 4</td>
<td>Truncated protein</td>
<td>4 (Chinese)</td>
<td>Wang et al. (1990b)</td>
</tr>
<tr>
<td>(28) Mis-sense</td>
<td>G→A$^b$ exon 7</td>
<td>Arg$^{243}$→Gln$^a$</td>
<td>4 (Chinese)</td>
<td>Wang et al. (1990b)</td>
</tr>
<tr>
<td>(29) Non-sense</td>
<td>G→A$^b$ exon 10</td>
<td>Trp$^{245}$→Ter$^a$</td>
<td>4 (Chinese)</td>
<td>Wang et al. (1990b)</td>
</tr>
<tr>
<td>(30) Non-sense</td>
<td>C→A or G?</td>
<td>Tyr$^{356}$→Ter$^a$</td>
<td>4 (Chinese)</td>
<td>Wang et al. (1990b)</td>
</tr>
<tr>
<td>(31) Deletion</td>
<td>exon 3</td>
<td>Ile$^{91}$ deleted</td>
<td>2 (Portuguese)</td>
<td>Caillaud et al. (1990)</td>
</tr>
</tbody>
</table>

$a$: Designates the study of the mutation by expression analysis

$b$: Indicates that the mutation involves a CpG dinucleotide

$c$: The abbreviation NA refers to the fact the mutation has not been assigned to a known haplotype due to the alteration of RFLP sites and/or the resulting DNA fragments by the mutation
ler et al. 1987a) and was also exclusively associated with the classic disease state, the Woo group proposed that a single mutation was likely to be responsible for PKU in patients carrying this mutant allele. The same reasoning was also applied to the mutant haplotype 2 allele, which shows both a similar distribution pattern and a clinical phenotype association. The Woo group subsequently characterized a mutation in the splice donor site of intron 12 (SPI2, DiLella et al. 1986b), which is associated with the Danish mutant haplotype 3 allele, and delineated a mis-sense mutation in exon 12 (Arg<sup>408</sup>→Trp, Di Lella et al. 1987) associated with the Danish haplotype 2 PKU allele. Based on the haplotype 2 (Arg<sup>408</sup>→Trp) and haplotype 3 (SPI2) mutations, it is possible to identify mutant PAH genes in 69% of East German PKU patients, 58% of Danish patients, declining to 42% for West German patients. It also permits the detection of about 1.5%, rather than 3%, of Turkish PKU alleles, since the single mutant haplotype 3 allele detected does not bear the SPI2 mutation.

At the present time, a total of 31 different mutations in the human PAH gene have been reported (see Table 1). Sixteen of the mutations listed in this table (numbers 1-4, 7-9, 14-19, and 28-30) have also been studied at the expression level (DiLella et al. 1987; Marvit et al. 1987; Lichter-Konecki et al. 1988a; Okano et al. 1990a, c; Wang et al. 1990b). The three larger deletions in the PAH gene (Table 1, numbers 20-22) have been localized to specific regions of the human PAH gene, although the greatly altered RFLP patterns accompanying these mutations preclude their assignment to known haplotypes.

While most of the mutations reported to date are relatively infrequent, six PKU mutations associated with haplotype 1 through 4 (Table 1; numbers 1, 2, 7, 8, 10, 19) are frequent among Caucasians. These DNA alterations are all single base transitions resulting either in amino acid substitutions or aberrant mRNA splicing.

To determine the frequency and distribution of these six most frequent, and the two additional less frequent (Table 1, nos. 9 and 16) mutations associated with haplotypes 1 through 4 in different Caucasian populations (Woo et al. 1991), we performed allele-specific oligonucleotide (ASO) hybridization on polymerase chain reaction (PCR)-amplified DNA (Saiki et al. 1985; reviewed by Vosberg 1989). A total of 132 PKU families from three different populations (99 German, 24 Turkish, and 9 Italian families), each harboring at least one haplotype 1, 2, 3, or 4 allele, were investigated using mutation-specific oligonucleotides as probes. The results of this investigation are presented in Tables 2 and 3. The haplotype 1 mutation in exon 7 (Arg<sup>201</sup> to Gln) was detected in 31% and 40% of the German and Turkish mutant haplotype 1 alleles, respectively, as against 72% of the mutant haplotype 1 alleles in Switzerland and 25% of mutant haplotype 1 alleles in Europe (Okano et al. 1990a).

The exon 12 mutation (Arg<sup>408</sup>→Trp) was found to be in close linkage with mutant haplotype 2 alleles in the three populations investigated, although it is very rare in the Turkish and Italian populations. Independent studies by several groups have also observed nearly complete linkage disequilibrium between the Arg<sup>408</sup>→Trp mis-sense mutation and PKU haplotype 2 alleles in the Danish (DiLella et al. 1987), German (Lichter-Konecki et al. 1988b); French (Rey et al. 1988), Swiss (Sullivan et al. 1989), Scottish (Sullivan et al. 1989) and Polish (Jurealska et al. 1991; Zyguńska et al. 1991) populations. Four studies have detected a lack of association specificity between the Arg<sup>408</sup> to Trp PKU mis-sense mutation and PAH-DNA haplotype 2. In the French study (Rey et al. 1988), all patients with a mutant haplotype 2 allele

### Table 2. Frequency of mutations, and their association with DNA haplotypes, in 123 PKU families of two populations. Each family possesses at least one mutant haplotype 1, 2, 3 or 4 allele. This study represents the analysis of 99 German and 24 Turkish PKU kindreds.

<table>
<thead>
<tr>
<th>Mutation</th>
<th>German PKU alleles</th>
<th>Turkish PKU alleles</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>H1</td>
<td>H2</td>
</tr>
<tr>
<td>Arg&lt;sup&gt;201&lt;/sup&gt;→Gln</td>
<td>12</td>
<td>-</td>
</tr>
<tr>
<td>Pro&lt;sup&gt;201&lt;/sup&gt;→Leu</td>
<td>7</td>
<td>-</td>
</tr>
<tr>
<td>Arg&lt;sup&gt;408&lt;/sup&gt;→Trp</td>
<td>53</td>
<td>3</td>
</tr>
<tr>
<td>SPI2</td>
<td>32</td>
<td>-</td>
</tr>
<tr>
<td>Arg&lt;sup&gt;408&lt;/sup&gt;→Gln</td>
<td>-</td>
<td>12</td>
</tr>
<tr>
<td>Tyr&lt;sup&gt;411&lt;/sup&gt;→Cys</td>
<td>-</td>
<td>9</td>
</tr>
<tr>
<td>Arg&lt;sup&gt;413&lt;/sup&gt;→Ter</td>
<td>-</td>
<td>3</td>
</tr>
<tr>
<td>Leu&lt;sup&gt;46&lt;/sup&gt;→Ser</td>
<td>-</td>
<td>3</td>
</tr>
<tr>
<td>N&lt;sup&gt;a&lt;/sup&gt;</td>
<td>40</td>
<td>53</td>
</tr>
<tr>
<td>F&lt;sup&gt;b&lt;/sup&gt; (%)</td>
<td>20%</td>
<td>26%</td>
</tr>
<tr>
<td>Af (%)</td>
<td>48%</td>
<td>100%</td>
</tr>
<tr>
<td>TP (%)</td>
<td>64%</td>
<td>34%</td>
</tr>
</tbody>
</table>

<sup>a</sup> N is the total number of mutant alleles investigated in each population

<sup>b</sup> F refers to the frequency (in %) of the mutant allele in the population

<sup>c</sup> Af is the percentage of PKU alleles possessing characterized PAH mutations

<sup>d</sup> TP is the total percentage of all PKU alleles investigated, found to possess identified mutations in the PAH gene

### Table 3. Frequency of mutations, and their association with DNA haplotypes, in nine Italian PKU families. At least one mutant haplotype 1, 2, 3, or 4 allele was possessed by each kindred.

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Italian PKU alleles</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>H1</td>
</tr>
<tr>
<td>Arg&lt;sup&gt;201&lt;/sup&gt;→Gln</td>
<td>1</td>
</tr>
<tr>
<td>Pro&lt;sup&gt;201&lt;/sup&gt;→Leu</td>
<td>1</td>
</tr>
<tr>
<td>Arg&lt;sup&gt;408&lt;/sup&gt;→Trp</td>
<td>-</td>
</tr>
<tr>
<td>SPI2</td>
<td>-</td>
</tr>
<tr>
<td>Arg&lt;sup&gt;408&lt;/sup&gt;→Gln</td>
<td>-</td>
</tr>
<tr>
<td>Tyr&lt;sup&gt;411&lt;/sup&gt;→Cys</td>
<td>-</td>
</tr>
<tr>
<td>Arg&lt;sup&gt;413&lt;/sup&gt;→Ter</td>
<td>-</td>
</tr>
<tr>
<td>Leu&lt;sup&gt;46&lt;/sup&gt;→Ser</td>
<td>-</td>
</tr>
<tr>
<td>N&lt;sup&gt;a&lt;/sup&gt;</td>
<td>7</td>
</tr>
<tr>
<td>F&lt;sup&gt;b&lt;/sup&gt; (%)</td>
<td>41%</td>
</tr>
</tbody>
</table>

<sup>a</sup> N is the total number of mutant alleles investigated in each respective population

<sup>b</sup> F refers to the frequency (in %) of the mutant allele in the population

Through the analysis of these tables, it becomes evident that the distribution of mutations and haplotypes varies significantly among different populations. This variability underscores the importance of genetic heterogeneity in the study of PKU and highlights the need for further research to better understand the role of genetic and environmental factors in the development of this disease.
and the phenotype of classic PKU carried the Arg⁴⁰⁸ to Trp mutation at this allele, while patients with less severe phenotypes did not. Whether these patients not exhibiting the expected association between mutation and haplotype were of European or North African ancestry was not indicated in the publication. Only two of four Italian PKU haplotype 2 alleles were found to harbor the Arg⁴⁰⁸→Trp mis-sense mutation (Dianzani et al. 1990b). John et al. (1990) have found this mutation in exon 12 to occur on haplotype 1 alleles in two French-Canadian PKU families. Recently, a Polish study (Zygulska et al. 1991) has reported the detection of the Arg⁴⁰⁸→Trp mutation on a single haplotype 5 PKU allele. This case may represent a new mutation, since a minimum of one recombinational and one mutational event would be required to transfer the Arg⁴⁰⁸ to Trp mis-sense mutation from a PKU haplotype 2 allele to a haplotype 5 allele. The intron 12 splice site mutation (SP12; DiLella et al. 1986b) has been found to be associated with all mutant haplotype 3 alleles analyzed in Denmark (DiLella et al. 1986b), France (Rey et al. 1988), Switzerland (Sullivan et al. 1989) and Scotland (Sullivan et al. 1989). In each of the following populations a single mutant haplotype 3 allele has been detected which has not been associated with the intron 12 splice site mutation: German (Aulehla-Scholz et al. 1988), Swedish (Svensson et al. 1991), Italian (Dianzani et al. 1990b) and Turkish (Konecki et al. 1991). The only PKU haplotype 3 allele analyzed in the Turkish population was associated with the Leu⁴⁸→Ser mutation, which was otherwise associated only with mutant haplotype 4 alleles (Konecki et al. 1991). Whether this mutation will be detected among other haplotype 3 alleles in the Mediterranean area, or whether it appeared on a PKU haplotype 3 background as the result of a recombinational or gene conversion event remains to be determined. This mutation (Leu⁴⁸ to Ser) in PAH exon 2 has since been detected on 7% of German, 33% of Turkish, and two of five Italian mutant haplotype 4 alleles. The exon 5 mis-sense mutation (Arg⁰⁵⁸ to Gln) was found to be associated with 27% of the German, 21% of the Turkish, and one of five Italian haplotype 4 alleles, compared with 33% of the mutant haplotype 4 alleles in the Swiss population and 40% of mutant haplotype 4 alleles in Europe (Okano et al. 1990a).

Among 12 mutant German haplotype 7 alleles, 2 show a loss of the BamHI site in exon 7 of the PAH gene. This was also observed in four of five mutant haplotype 7 alleles in the Swedish population (Svensson et al. 1990) and 82% of PKU haplotype 7 alleles in Norway, which comprise about 20% of the mutant PAH alleles in the Norwegian population (Apold et al. 1990). In the Swedish and Norwegian populations, the loss of the BamHI site has been shown to be the result of a single base transition generating a termination codon (Arg⁴⁵³→Ter; Svensson et al. 1990). This non-sense mutation, resulting in an inactive and severely truncated form of the PAH enzyme, is responsible for PKU in these patients. Since only 2 of 12 German, 4 of 5 Swedish, and 14 of 17 Norwegian mutant haplotype 7 alleles show the loss of the BamHI restriction endonuclease recognition site, at least one more mutation must have arisen on the background of this allele, which is the third most frequent normal allele in most populations.

Thus far, five different mutations have been detected on haplotypic backgrounds differing from those on which the mutations were originally characterized: (1) the exon 12 haplotype 2 mutation (Arg⁴⁰⁸ to Trp) on a haplotype 1 background in French Canadians (John et al. 1990), on a haplotype 1 background in French Canadians (John et al. 1990), on a haplotype 4 background in Chinese (Tsai et al. 1990), and on a haplotype 5 background in a Polish individual (Zygulska et al. 1991); (2) the exon 7 haplotype 38 mutation (Glu²⁶⁰ to Lys) on a haplotype 4 background (Abadie et al. 1989) and a haplotype 1 background (Okano et al. 1990b); (3) the exon 7 Pro²⁰⁴ to Leu mutation on haplotype 1 and 4 backgrounds (Okano et al. 1989); (4) the exon Leu²⁸⁰→Pro mutation on haplotype 1, 10, and 7 backgrounds (Lichter-Konecki et al. 1988a; Riess et al. 1988; Hofman et al. 1988); and (5) the exon 2 Leu⁴⁸→Ser mutation on haplotype 3 and 4 backgrounds (Konecki et al. 1991). It seems quite remarkable that the exon 12 mis-sense mutation (Arg⁴⁰⁸ to Trp) has now been detected on four haplotypic backgrounds and the exon 7 Glu²⁶⁰ to Lys mutation on three haplotypic backgrounds. While recombinational or gene conversion events must be considered as a means of transferring such mutations from one allele to another, these mechanisms appear relatively unlikely for these two mutations (Arg⁴⁰⁸→Trp and Glu²⁶⁰→Lys). Only in the case of the exon 12 mis-sense mutation being transferred from a haplotype 2 allele to a haplotype 1 allele (or vice versa) would a single recombinational event be sufficient to explain such a transfer. A more likely explanation concerns the nature of the bases involved in these two mutations. In both instances (the Arg⁴⁰⁸→Trp and the Glu²⁶⁰→Lys mis-sense mutations), the substituted base is a member of a CpG dinucleotide, which has been shown to be highly mutable, providing a "hot spot" for mutation (Cooper and Youossoufian 1988; Abadie et al. 1989). Thus far, ten different mutations (Table 1, nos. 2, 4, 5, 7-9, 15, 16, 18, and 28) involving nine CpG dinucleotides have been identified. It remains to be determined exactly how many of the 22 CpG dinucleotides contained in the transcribed region encoding the PAH enzyme have incurred mutational alteration.

Conclusions

Association between DNA haplotype and mutation

In summary, the initial hypothesis of association between mutations and DNA haplotypes at the PAH locus remains a general rule, the few exceptions having been mentioned in the preceding paragraph. The distributions of DNA haplotypes in different populations have been observed to vary considerably. Mutant haplotype 2 and 3 alleles are frequent among European populations north of the Alps and there are specific mutations associated with them. By contrast these alleles, and their associated mutations, are of little significance in European populations south of the Alps. A different haplotype 2
mutation (Met to Val) has been observed among French Canadian PKU patients (John et al. 1990). Since haplotypes 1 and 4 are found in essentially all populations investigated, with approximately equal frequencies among normal and PKU chromosomes, these haplotypes are considered to be the most ancient alleles. Also, due to their relatively high frequencies, multiple mutations are assumed to have occurred on haplotype 1 and 4 backgrounds. Indeed, five different mutations have already been identified in Caucasian haplotype 4 alleles (Arg to Gln, Dworniczak et al. 1989; Okano et al. 1990a; Glu to Gly, Abadie et al. 1989; Pro to Leu, Okano et al. 1989; Arg to Ser, Wang et al. 1990a; Leu to Ser, Konecki et al. 1991) and four different mutations in haplotype 1 alleles (Leu to Pro, Lichter-Konecki et al. 1988a; Glu to Lys, Okano et al. 1990b; Arg to Gln, Okano et al. 1990a; Pro to Leu, Okano et al. 1989; Arg to Trp, Okano et al. 1989). These nine mutations have been observed with different frequencies in different Caucasian ethnic groups, suggesting population-specific patterns. Another nine mutations (Table 1, nos. 5, 17, 18, 22, and 26–30), most of which are tightly linked to PAH-DNA haplotype 4 alleles, are observed only in Oriental populations, supporting the hypothesis of the recent occurrence of PKU mutations on haploptic backgrounds after divergence of the races (Levy 1989). The only DNA alterations within the region coding for the PAH enzyme which have been detected in Orientals as well as Caucasians are the silent mutations (Gln to Gln and Val to Val, Lichter-Konecki et al. 1990) residing in exon 6 and 7, respectively. Our laboratory has found both of these same-sense mutations to be associated with mutant as well as normal haplotype 4 alleles, whereas the Gln to Gln silent mutation was also associated with normal and mutant haplotype 3 and 7 alleles in all populations investigated (German, Italian, Turkish, Kuwaiti, and Japanese).

Therefore, with few exceptions, distinct mutations are associated with specific DNA haplotypes, although more than one mutation may be associated with a particular haplotype. This is especially true with regard to the most frequent haplotypes in different populations. Mutations, and alleles, of the PAH gene also show population-specific distributions.

Origins of the PAH mutations

Obviously, there must have been different origins, throughout the world, for the various mutations in the PAH genes which result in disorders of phenylalanine hydroxylase deficiency. Haplotypes 2 and 3, and the mutations associated with them, seem to have occurred relatively recently in European populations north of the Alps, with the highest frequency of mutant haplotype 3 alleles reported for Denmark and of mutant haplotype 2 alleles for Czechoslovakia. Therefore, one may speculate that these might be the populations in which the respective alleles arose and from which they may have spread. Such a theory is not only supported by the gradual decline in the frequency of the haplotype 2 allele from the east to the west of Europe, but also with the observation of an increased incidence of PKU in northwestern Germany after World War II, resulting from the migration of northeastern German families into this area (Flatz et al. 1984). Evaluation of a substantial collection of European PAH haplotype analysis data (Daiger et al. 1989a) has failed to confirm the earlier suggestion for a Celtic origin of PKU in Europe (Carter and Woolf 1961).

The different haplotype 4 mutations obviously have different origins as exemplified by the varying frequencies with which these mutations are observed in different populations. For example, the exon 2 mutation (Leu to Ser; Konecki et al. 1991) is much more frequent in Turkey than among the other Caucasian populations investigated, while the mutation in exon 5 (Arg to Gln) is equally frequent in most Caucasians populations. A third mutation (Arg to Ser in exon 7, Wang et al. 1990a), associated with Caucasian haplotype 4 alleles, was originally only observed once each in the populations of Hungary and Czechoslovakia, and was subsequently found to occur with similarly low frequencies in our studies of the German, Turkish, and Italian populations. This may, in future, be found to represent a significant mutation in an ethnic group that remains to be analyzed. The haplotype 1 mutation (Arg to Gln), originally characterized in a Swiss individual, was detected in 72% of the mutant haplotype 1 alleles in Switzerland; however, in other European and in Eurasian populations it has been found in lower frequencies (25–40%) of mutant haplotype 1 alleles (Okano et al. 1989; 1990a; this publication). Founder effect has clearly participated in the spread of the deletion of exon 3 among the Yemenite Jewish population (Avigad et al. 1990) and the mutation involving the initiation codon in exon 1 (Met to Val) among French Canadians (John et al. 1989).

Explanations for the high frequency of mutant PAH alleles

High mutation rate, hitchhiking of PKU alleles through selection at a nearby locus (interferon Y), founder effect, genetic drift, and natural selection for PKU heterozygotes have all been considered as possible mechanisms for the observed high incidence of PKU (Scriver 1986; Kidd 1987; Scriver et al. 1989). The discovery of two mutations (SP12 and Arg to Trp) accounting for more than 50% of the PKU alleles in Denmark (and 42% of those in West Germany) appears to exclude the theory of a high mutation rate at the PAH locus as a cause for the high frequency of PKU. Random genetic drift is excluded, as a general rule, because it would have to affect several different PKU alleles simultaneously in a similar manner.

While “founder effect” certainly seems to play a role in the spread of PKU mutations in isolated populations (e.g., in Yemenite Jews and French Canadians), natural selection as the result of a compensating advantage appears to be a more plausible mechanism for the high allele frequencies and the observed high incidence of PKU in many different modern populations (Scriver 1986; Kidd 1987; Woo 1989). The exact nature of such a compensating heterozygote advantage has been a topic of speculation for many years (Saugstad 1976; Woolf 1976,
1986; Vogel 1984; Trefz et al. 1989). A recent hypothesis concerns the increased viability of the fetus, afforded by modest hyperphenylalaninemia in the pregnant heterozygote against exposure to ochratoxin A (Woo 1989). This compound is a known ubiquitous mycotoxin abortifacient.

**Phenotypic heterogeneity of the hyperphenylalaninemas**

Taking into consideration the above discussion concerning haplotypes and mutations at the human PAH locus, there are now several explanations for the phenotypic heterogeneity of PAH deficiency at the molecular level. First, it has been shown that different mutations at the PAH locus result in classic PKU (DiLella et al. 1986b; Avigad et al. 1990) or its milder variants (Lyonnet et al. 1989; Okano et al. 1990a, c; Konecki et al. 1991). Most DNA alterations in the PAH gene that have thus far been shown to be associated with phenylalanine hydroxylase deficiency involve single base substitutions and deletions. Therefore, it is clear that no unique DNA alteration of the PAH gene is responsible for this group of metabolic disorders. Second, the majority of Caucasian PKU patients so far analyzed possess two different PAH haplotypes and can be considered haplotypic heterozygotes. Such patients are also compound heterozygotes from the standpoint of mutations, since different haplotypes were shown to be associated with different mutations. After the characterization of the Arg<sup>408</sup>→Trp and SP12 mutations (associated with PKU haplotypes 2 and 3, respectively), it became evident that homozygous or exclusively heterozygous (i.e., Arg<sup>408</sup>→Trp/SP12 compound heterozygotes) patients manifest a classic PKU phenotype (Gütüler et al. 1987b; Herrmann et al. 1988; Lichter-Konecki et al. 1988b). An even more complex picture emerges from the detection of multiple mutations associated with the mutant haplotype 1 and 4 alleles.

The investigation of 138 of our patients (99 German, 24 Turkish, 9 Italian, and 6 others) possessing at least one haplotype 1–4 allele for six relatively frequent (Table 1; nos. 1, 2, 7, 8, 10, and 19) and two less common (Table 1, nos. 9 and 16) mutations allowed complete genotype analysis with regard to the mutations in the PAH genes of 59 individuals. For 52 of these patients, sufficient clinical data (i.e., pretreatment plasma phenylalanine levels and results of protein loading studies) had been collected to permit a correlation of the genotypes with the clinical phenotypes (Okano et al. 1991). The results of this correlation are displayed in Table 4.

<table>
<thead>
<tr>
<th>Mutation in PAH gene (1st allele)</th>
<th>Mutation in PHA gene (2nd allele)</th>
<th>RFLP haplotypes</th>
<th>Residual activity&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Clinical phenotype&lt;sup&gt;b&lt;/sup&gt;</th>
<th>No. of patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arg&lt;sup&gt;408&lt;/sup&gt;→Trp</td>
<td>Arg&lt;sup&gt;408&lt;/sup&gt;→Trp</td>
<td>2/2</td>
<td>0/0</td>
<td>I</td>
<td>7</td>
</tr>
<tr>
<td>Pro&lt;sup&gt;381&lt;/sup&gt;→Leu</td>
<td></td>
<td>2/0</td>
<td>0/0</td>
<td>I</td>
<td>3</td>
</tr>
<tr>
<td>Arg&lt;sup&gt;631&lt;/sup&gt;→Gln</td>
<td></td>
<td>2/1</td>
<td>0/30</td>
<td>I</td>
<td>3</td>
</tr>
<tr>
<td>Arg&lt;sup&gt;158&lt;/sup&gt;→Gln</td>
<td></td>
<td>2/4</td>
<td>0/10</td>
<td>I</td>
<td>4</td>
</tr>
<tr>
<td>SP12&lt;sup&gt;c&lt;/sup&gt;</td>
<td>SP12</td>
<td>3/3</td>
<td>0/0</td>
<td>I</td>
<td>3</td>
</tr>
<tr>
<td>Arg&lt;sup&gt;408&lt;/sup&gt;→Trp</td>
<td></td>
<td>3/2</td>
<td>0/0</td>
<td>I</td>
<td>8</td>
</tr>
<tr>
<td>Arg&lt;sup&gt;631&lt;/sup&gt;→Gln</td>
<td></td>
<td>3/1</td>
<td>0/30</td>
<td>I</td>
<td>1</td>
</tr>
<tr>
<td>Leu&lt;sup&gt;38&lt;/sup&gt;→Ser</td>
<td></td>
<td>3/4</td>
<td>0/n.d.</td>
<td>I</td>
<td>1</td>
</tr>
<tr>
<td>Glu&lt;sup&gt;221&lt;/sup&gt;→Gly</td>
<td></td>
<td>3/4</td>
<td>ND/ND</td>
<td>I</td>
<td>1</td>
</tr>
<tr>
<td>Arg&lt;sup&gt;158&lt;/sup&gt;→Gln</td>
<td></td>
<td>4/4</td>
<td>ND/10</td>
<td>I</td>
<td>1</td>
</tr>
<tr>
<td>Arg&lt;sup&gt;158&lt;/sup&gt;→Gln</td>
<td></td>
<td>4/4</td>
<td>10/10</td>
<td>I</td>
<td>4</td>
</tr>
<tr>
<td>Arg&lt;sup&gt;261&lt;/sup&gt;→Cys</td>
<td>Arg&lt;sup&gt;631&lt;/sup&gt;→Gln</td>
<td>1/1</td>
<td>30/30</td>
<td>II</td>
<td>4</td>
</tr>
<tr>
<td>Tyr&lt;sup&gt;144&lt;/sup&gt;→Cys</td>
<td>Tyr&lt;sup&gt;144&lt;/sup&gt;→Cys</td>
<td>1/4</td>
<td>30/50</td>
<td>II</td>
<td>2</td>
</tr>
<tr>
<td>Pro&lt;sup&gt;381&lt;/sup&gt;→Cys</td>
<td>Tyr&lt;sup&gt;144&lt;/sup&gt;→Cys</td>
<td>1/4</td>
<td>0/50</td>
<td>II</td>
<td>1</td>
</tr>
<tr>
<td>Tyr&lt;sup&gt;144&lt;/sup&gt;→Cys</td>
<td>Arg&lt;sup&gt;408&lt;/sup&gt;→Trp</td>
<td>4/2</td>
<td>50/0</td>
<td>I-II</td>
<td>2</td>
</tr>
<tr>
<td>Arg&lt;sup&gt;408&lt;/sup&gt;→Trp</td>
<td>Arg&lt;sup&gt;408&lt;/sup&gt;→Trp</td>
<td>4/2</td>
<td>50/0</td>
<td>II</td>
<td>3</td>
</tr>
<tr>
<td>SP12</td>
<td></td>
<td>4/3</td>
<td>50/0</td>
<td>II</td>
<td>2</td>
</tr>
<tr>
<td>Leu&lt;sup&gt;38&lt;/sup&gt;→Ser</td>
<td>Leu&lt;sup&gt;38&lt;/sup&gt;→Ser</td>
<td>4/4</td>
<td>ND</td>
<td>II</td>
<td>2</td>
</tr>
</tbody>
</table>

<sup>a</sup> Residual PAH enzyme activity determined as described by Okano et al. 1990a, c
<sup>b</sup> I, Classic PKU; II, mild PKU; III, hyperphenylalaninemia (HPA)
<sup>c</sup> SP12 refers to the mutation in the splice donor site of intron 12
<sup>d</sup> The Leu<sup>38</sup>→Ser is otherwise only associated with haplotype 4 alleles

The intron 12 splice site mutation is almost exclusively associated with the classic phenotype and shows complete concordance with mutant PAH haplotype 3 alleles. Only in those patients possessing a second mutant allele coding for an enzyme with very high residual PAH activity i.e., Tyr<sup>144</sup> to Cys; Okano et al. 1990) does the effect of a single haplotype 3 allele bearing the SP12 mutation not result in a severe PKU phenotype. The exon 12 (Arg<sup>408</sup>→Trp) mis-sense mutation has been shown to be associated with the classic phenotype either in patients homozygous for the mutation or in those whose other mutant PAH allele carries the intron 12 splice site mutation. While known to occur in tight linkage with PKU haplotype 2 alleles (DiLella et al. 1987), this mutation has also been detected in two other mutant PAH haplotypes (John et al. 1990; Tsai et al. 1990). In combination with a mutant allele coding for an altered PAH enzyme with high residual activity, the Arg<sup>408</sup>→Trp mutation has also been observed in association with a mild phenylalanine hydroxylase deficiency phenotype (Lichter-Konecki et al. 1988b, 1989b). Both of these mutations affecting PAH exon 12 (SP12 and Arg<sup>408</sup>→Trp) result in no detectable residual enzyme activity in the heterogeneous eukaryotic expression systems previously described (DiLella et al. 1987; Marvit et al. 1987; Okano et al. 1990a, c).

The exon 7 (Arg<sup>261</sup>→Gln) mutation associated with haplotype 1 (Abadie et al. 1989) was originally reported not to result in any decrease of PAH enzyme activity (Okano et al. 1990a). However, subsequent expression analyses have revealed a residual activity, of approximately 30% normal, in the assay system (Okano et al. 1990c). Patients homozygous for this mis-sense mutation exhibit the phenotype of mild PKU. The base substitu-
tion in codon 158, located in PAH exon 5, is associated with haplotype 4 alleles. The modified PAH enzyme resulting in the substitution of Gin for Arg yielded a PAH enzyme possessing approximately 10% residual activity in the expression system used (Okano et al. 1990a, c) and was associated with classic PKU in patients homozygous for it. The mis-sense mutation Tyr^{324}→Cys, identified in PAH exon 12 and associated with haplotype 4 alleles, has been shown to yield the highest residual enzyme activity (50% of normal in the assay system) of all DNA alterations so far investigated at the expression level (Okano et al. 1990c). This mutation is associated with the milder phenotypes even when the second allele harbors a mutation resulting in no detectable residual enzyme activity. The haplotype 4 mutation in exon 2 (Leu^{46}→Ser^{48}) is associated with mild PKU in these patients who are homozygous for it (Konecki et al. 1991).

**Perspectives on the direct detection of PKU mutations**

With the advent of the polymerase chain reaction (PCR) the number of groups involved in the study of mutations and polymorphisms of the PAH gene has increased. Accordingly, the number of DNA alterations reported has also grown. Only four of the fully characterized mutations listed in Table 1 (nos. 1–4) were identified without the use of in vitro DNA amplification. It can be expected that within the near future the human PAH gene will be sequenced in its entirety, although targeted sequencing will probably be directed at regions containing the polymorphic restriction sites currently used for haplotype analysis. Knowledge of the sequences flanking these sites would allow implementation of PCR amplification and ASO hybridization for haplotype analysis and greatly accelerate this laborious process, which forms the basis for the study of mutations at the PKU locus.

Most of the DNA alterations so far reported at this genetic locus represent mutations abolishing PAH enzyme activity and resulting in the most severe form of phenylalanine hydroxylase deficiency. Undoubtedly a phase will be entered in which the DNA changes identified will produce more subtle alterations in the enzyme, necessitating the detailed study of the kinetics of mutated enzymes. Fortunately, systems currently exist and are already being implemented for this purpose, which should provide further insight into the regulation and expression of the human PAH gene and its product.

As shown in Table 2, about 64% and 34% of the mutations in the German and Turkish populations, respectively, can be detected through the use of ASO probes for eight of the most frequent mutations tested to date. Too few Italian PKU families have been studied to assess the effectiveness of these probes for screening purposes. In most populations it is not yet possible to detect more than 50% of the DNA alterations responsible for disorders due to PAH deficiency. Obviously, many more mutations remain to be characterized for better understanding of the phenotypic heterogeneity of disorders caused by a deficiency in the PAH enzyme, as well as a more thorough knowledge of the regulatory and functional domains of that enzyme. Such knowledge should result in better patient care by improving the predictability of the clinical course (and outcome) of the disease and may allow for the development of new modes of treatment (e.g., enzyme therapy).

**Acknowledgements.** The authors wish to thank Professor F. Vogel for helpful discussions and for critical reading of the manuscript, as well as Dr. F. K. Trefz and M. Schlatter for technical assistance. Professor S. L. C. Woo is also thanked for providing the recombinant pHPAH247 and for stimulating discussions. This work was supported by grants from the Deutsche Forschungsgemeinschaft: Tr.175/2-1 (to Dr. F. K. Trefz and Professor Dr. A. Sippel), Li.375/1-1/1-2 and Li.375/2-1/2-2 (to U.-L.-K.). Support from the Fritz Thyssen Foundation is also acknowledged (grant: 1990/1/51, to Professor Dr. H. J. Bremer).

**References**


Daiger SP, Chakraborty R, Reed L, Fekete G, Schuler D, Berensti
van S, Woo SLC (1989) Polymorphic DNA haplotypes at the ph
ylenalanine hydroxylase (PAH) locus in European families with
phenylketonuria (PKU) Am J Hum Genet 45:310–318
Daiger SP, Reed L, Huang SS, Zeng YT, Wang T, Lo WHY, Okano
Y, Hase Y, Fukuda Y, Oura T, Tada K, Woo SLC (1989b) Po
lymorphic DNA haplotypes at the phenylalanine hydroxylase
(PAH) locus in Asian families with phenylketonuria (PKU)
Am J Hum Genet 45:319–324
Dianzani I, Camaschella C, Saglio G, Ferreri GB, Romeo G, De
vito M, Romano C, Cerone R, Giovannini M, Riva E, Trefz
FK, Lichter-Konecki U, Woo SLC (1990a) Haplotype distribu
tion and molecular defects of PKU in Italy. J Inher Metab Dis
13:292–294
Dianzani I, Devito M, Camaschella C, Saglio G, Ferreri GB,
Cerone R, Romano C, Romeo G, Giovannini M, Riva E, An
geneydt F, Trefz FK, Okano Y, Woo SLC (1990b) Haplo
type distribution and molecular defects of the phenylalanine
hydroxylase locus in Italy. Hum Genet 86:69–72
DiLella AG, Kwok SCM, Ledley FD, Marvit J, Woo SLC (1986a)
Molecular structure and polymorphic map of the human phenyl
DiLella AG, Marvit J, Lidsky AS, Gütter F, Woo SLC (1986b)
Tight linkage between a splicing mutation and a specific DNA
cid substitution involved in phenylketonuria is in linkage dis
ria: detection of a frequent haplotype 4 allele mutation. Hum
Genet 84:95–96
hydroxygense: silent mutation uncovers evolutionary origin
phenylketonuria in the North German population. Hum Genet
65:396–399
Fölling A (1934) Über Ausscheidung von Phenylbrenztraubensäure
in den Harn als Stoffwechselanomalie in Verbindung mit Imbe
Gütter F (1984) Phenylketonuria: 50 years since Fölling’s discov
ery and still expanding our clinical and biochemical knowledge.
Acta Paediatr Scand 73:705–716
Gütter F, DiLella AG, Ledley FD, Lidsky AS, Kwok SCM, Mar
Eur J Pediatr 146 [Suppl 1]:A5–A11
Gütter F, Ledley FD, Lidsky AS, DiLella AG, Sullivan SE, Woo
SLC (1987b) Correlation between polymorphic DNA haplo
types at phenylalanine hydroxylase locus and clinical pheno
tecting phenylketonuria in large populations of newborn in
fants. Pediatrics 32:338–343
Haplotyp analysis of classical and mild phenotype of phenyl
ketonuria in the German Democratic Republic. Clin Genet 34:
176–180
gene haplotypes in Polynesians: evolutionary origins and ab
sence of alleles associated with severe phenylketonuria. Am J
Hum Genet 44:382–387
Hofman KJ, Antonarakis SE, Miissou-Tsangaraki S, Boehm
CD, Vallee D (1989) Phenylketonuria in the Greek population:
ahaplotype analysis of the phenylalanine hydroxylase gene
256
GTA(Val)→GTT(Val), a new silent mutation found in the
Chinese. Hum Genet 86:305–306
Jaruzelska J, Henriksen KF, Gütter F, Riess O, Borki K, Bln N,
Slomski R (1991) The codon 408 mutation associated with haplo
type 2 is predominant in Polish families with phenylketonu
ria. Hum Genet 86:247–250
Jervis GA (1947) Studies of phenylpyruvic ophthalmia: the po
Jervis GA (1953) Phenylpyruvic ophthalmia deficiency of phenyl
John SWM, Rozen R, Laframboise R, Laberge C, Scriber CR
(1989) Novel PKU mutation on haplotype 2 in French-Canadi
John SWM, Rozen R, Scriber CR, Laframboise R, Laberge C
(1990) Recurrent mutation, gene conversion, or recombination
at the human phenylalanine hydroxylase locus: evidence in
French-Canadians and a catalog of mutations. Am J Hum Genet
46:970–974
Kaufman S (1957) The enzymatic conversion of phenylalanine to
Kaufman S (1958) Phenylalanine cofactor in phenylketonuria. Scie
ence 128:1506–1508
Neuropathol 2:1–132
Kazazian HH, Orkin H, Markham AF, Chapman CR, Yousouf
ian H, Waber PG (1984) Quantification of the close associ
ation between DNA haplotypes and specific f-thalassemia mu
283
The identification of two mis-sense mutations at the PAH gene
locus in a Turkish patient with phenylketonuria. Hum Genet
87:389–393
Kwok SCM, Ledley FD, DiLella AG, Robson KJH, Woo SLC
(1985) Nucleotide sequence of a full-length complementary
dNA clone and amino acid sequence of human phenylalanine
hydroxylase. Biochemistry 24:556–561
Ledley FD, Grennet HE, DiLella AG, Kwok SCM, Woo SLC
(1985) Gene transfer and expression of human phenylalanine
hydroxylase. Science 228:77–79
Levy HL (1989) Invited editorial: Molecular genetics of phenyl
Lichter-Konecki U, Konecki DS, DiLella AG, Brayton K, Marvit
J, Hahn TM, Trefz FK, Woo SLC (1988a) Phenylalanine hydro
xylase deficiency caused by a single base substitution in an exon
of the human phenylalanine hydroxylase gene. Biochemistry
27:2881–2885
Lichter-Konecki U, Schlotter M, Konecki DS, Labeit S, Woo
SLC, Trefz FK (1988) Linkage disequilibrium between muta
tion and RFLP haplotype at the phenylalanine hydroxylase loc
us in the German population. Hum Genet 78:347–352
Lichter-Konecki U, Schlotter M, Yaylak C, Özgüc M, Coskun T,
Özalp I, Wendel U, Batzler U, Trefz FK, Konecki D (1989a)
DNA haplotype analysis at the phenylalanine hydroxylase loc
us in the Turkish population. Hum Genet 81:373–376
Lichter-Konecki U, Schlotter M, Trefz FK, Konecki D (1989b)
Direct detection of a major PKU mutation in the German popula
tion after DNA amplification. Eur J Pediatr 149:120–123
Identification of new mutation at the phenylalanine hydroxyl
ase gene locus. 5th International Congress of Inborn Errors of
Metabolism, Asilomar, abstract W 4.4
Lidsky AS, Law ML, Merce HG, Kao F-T, Rabin M, Ruddle FH,
Woo SLC (1985a) Regional mapping of the phenylalanine hydro
xylase gene and the phenylketonuria locus in the human geno
me. Proc Natl Acad Sci USA 82:6221–6225
Lidsky AS, Ledley FD, DiLella AG, Kwok SCM, Daiger SP,
Robson KJH, Woo SLC (1985b) Extensive restriction site poly
morphism at the human phenylalanine hydroxylase locus and
application in prenatal diagnosis of phenylketonuria. Am J
Hum Genet 37:619–624


and the phenotype of classic PKU carried the Arg^{408} to Trp mutation at this allele, while patients with less severe phenotypes did not. Whether these patients not exhibiting the expected association between mutation and haplotype were of European or North African ancestry was not indicated in the publication. Only two of four Italian PKU haplotype 2 alleles were found to harbor the Arg^{408}→Trp missence mutation (Dianzani et al. 1990b). John et al. (1990) have found this mutation in exon 12 to occur on haplotype 1 alleles in two French-Canadian PKU families. Recently, a Polish study (Zygulska et al. 1991) has reported the detection of the Arg^{408}→Trp mutation on a single haplotype 5 PKU allele. This case may represent a new mutation, since a minimum of one recombinational and one mutational event would be required to transfer the Arg^{408} to Trp missence mutation from a PKU haplotype 2 allele to a haplotype 5 allele. The intron 12 splice site mutation (SP12; DiLella et al. 1986b) has been found to be associated with all mutant haplotype 3 alleles analyzed in Denmark (DiLella et al. 1986b), France (Rey et al. 1988), Switzerland (Sullivan et al. 1989) and Scotland (Sullivan et al. 1989). In each of the following populations a single mutant haplotype 3 allele has been detected which has not been associated with the intron 12 splice site mutation: German (Aulehla-Scholz et al. 1988), Swedish (Svensson et al. 1991), Italian (Dianzani et al. 1990b) and Turkish (Konecki et al. 1991). The only PKU haplotype 3 allele analyzed in the Turkish population was associated with the Leu^{48}→Ser mutation, which was otherwise associated only with mutant haplotype 4 alleles (Konecki et al. 1991). Whether this mutation will be detected among other haplotype 3 alleles in the Mediterranean area, or whether it appeared on a PKU haplotype 3 background as the result of a recombinational or gene conversion event remains to be determined. This mutation (Leu^{48}→Ser) in PAH exon 2 has since been detected on 7% of German, 33% of Turkish, and two of five Italian mutant haplotype 4 alleles. The exon 5 mis-sense mutation (Arg^{219} to Gln) was found to be associated with 27% of the German, 21% of the Turkish, and one of five Italian haplotype 4 alleles, compared with 33% of the mutant haplotype 4 alleles in the Swiss population and 40% of mutant haplotype 4 alleles in Europe (Okano et al. 1990a).

Among 12 mutant German haplotype 7 alleles, 2 show a loss of the BamHI site in exon 7 of the PAH gene. This was also observed in four of five mutant haplotype 7 alleles in the Swedish population (Svensson et al. 1990) and 82% of PKU haplotype 7 alleles in Norway, which comprise about 20% of the mutant PAH alleles in the Norwegian population (Apold et al. 1990). In the Swedish and Norwegian populations, the loss of the BamHI site has been shown to be the result of a single base transition terminating a termination codon (Arg^{27}→Ter; Svenson et al. 1990). This non-sense mutation, resulting in an inactive and severely truncated form of the PAH enzyme, is responsible for PKU in these patients. Since only 2 of 12 German, 4 of 5 Swedish, and 14 of 17 Norwegian mutant haplotype 7 alleles show the loss of the BamHI restriction endonuclease recognition site, at least one more mutation must have arisen on the background of this allele, which is the third most frequent normal allele in most populations.

Thus far, five different mutations have been detected on haplotypic backgrounds differing from those on which the mutations were originally characterized: (1) the exon 12 haplotype 2 mutation (Arg^{408} to Trp) on a haplotype 1 background in French Canadians (John et al. 1990), on a haplotype 1 background in French Canadians (John et al. 1990), on a haplotype 4 background in Chinese (Tsai et al. 1990), and on a haplotype 5 background in a Polish individual (Zygulska et al. 1991); (2) the exon 7 haplotype 38 mutation (Glu^{269} to Lys) on a haplotype 4 background (Abadie et al. 1989) and a haplotype 1 background (Okano et al. 1990b); (3) the exon 7 Pro^{251} to Leu mutation on haplotype 1 and 4 backgrounds (Okano et al. 1989); (4) the exon Leu^{401}→Pro mutation on haplotype 1, 10, and 7 backgrounds (Lichter-Konecki et al. 1988a; Riess et al. 1988; Hofman et al. 1989); and (5) the exon 2 Leu^{56}→Ser mutation on haplotype 3 and 4 backgrounds (Konecki et al. 1991). It seems quite remarkable that the exon 12 missense mutation (Arg^{408} to Trp) has now been detected on four haplotypic backgrounds and the exon 7 Glu^{269} to Lys mutation on three haplotypic backgrounds. While recombinational or gene conversion events must be considered as a means of transferring such mutations from one allele to another, these mechanisms appear relatively unlikely for these two mutations (Arg^{408}→Trp and Glu^{269}→Lys). Only in the case of the exon 12 missense mutation being transferred from a haplotype 2 allele to a haplotype 1 allele (or vice versa) would a single recombinational event be sufficient to explain such a transfer. A more likely explanation concerns the nature of the bases involved in these two mutations. In both instances (the Arg^{408}→Trp and the Glu^{269}→Lys missense mutations), the substituted base is a member of a CpG dinucleotide, which has been shown to be highly mutable, providing a "hot spot" for mutation (Cooper and Vousvoyian 1988; Abadie et al. 1989). Thus far, ten different mutations (Table 1, nos. 2, 4, 5, 7–9, 15, 16, 18, and 28) involving nine CpG dinucleotides have been identified. It remains to be determined exactly how many of the 22 CpG dinucleotides contained in the transcribed region encoding the PAH enzyme have incurred mutational alteration.

Conclusions

Association between DNA haplotype and mutation

In summary, the initial hypothesis of association between mutations and DNA haplotypes at the PAH locus remains a general rule, the few exceptions having been mentioned in the preceding paragraph. The distributions of DNA haplotypes in different populations have been observed to vary considerably. Mutant haplotype 2 and 3 alleles are frequent among European populations north of the Alps and there are specific mutations associated with them. By contrast these alleles, and their associated mutations, are of little significance in European populations south of the Alps. A different haplotype 2