Am. J. Hum. Genet. 60:95–102, 1997

Analysis of the Phenylalanine Hydroxylase Gene in the Spanish Population: Mutation Profile and Association with Intragenic Polymorphic Markers

Belén Pérez, Lourdes R. Desviat, and Magdalena Ugarte

Centro de Biología Molecular "Severo Ochoa," CSIC-UAM, Madrid

Summary

The aim of this study was to characterize the phenylketonuria (PKU) alleles in the Spanish population, by both identifying the causative mutations and analyzing the RFLP haplotypes and the VNTR and short-tandem-repeat alleles associated with the phenylalanine hydroxylase (PAH) gene. We have investigated 129 independent mutant chromosomes, using denaturing gradient gel electrophoresis (DGGE) and direct sequencing. Ninety percent of the alleles were identified, and a total of 40 different mutations were detected. The mutational spectrum includes seven previously unreported mutations: P122Q, D129G, P147S, D151G, A165T, S196fs, and P407S. Seven mutations represent 43% of the Spanish PKU alleles, the most common being IVS10nt-11g→a (14.7%), I65T (8.5%), and V388M (6.2%). The remaining 33 mutations are rare. The mutation profile and relative frequencies are markedly different from those in northern Europe, also showing unique features compared with those in other, southern European populations. The association analysis with polymorphic markers in the PAH gene provides valuable information for population-genetic studies and investigation of the origins of the mutations. This study may serve as reference in the analysis of the contemporary distributions and frequencies of the PKU mutations in related populations, with particular relevance in Latin American countries.

Introduction

Hyperphenylalaninemia, caused by a deficiency of hepatic phenylalanine hydroxylase (PAH), is the most common inborn error of amino acid metabolism in Caucasians (Scriver et al. 1995). The PAH enzyme catalyzes the irreversible hydroxylation of phenylalanine to tyrosine. PAH deficiency results in accumulation of phenylalanine and secondary metabolites in physiological fluids and in untreated patients is manifested clinically as mental retardation. Dietary therapy based on a restricted phenylalanine intake prevents the neurological damage. This has led to the implementation of mass neonatal screening programs for PKU in Western countries.

The PAH gene was cloned in 1985 (Kwok et al. 1985), and, to date, >290 different mutations have been detected (PAH Mutation Analysis Consortium Database; Hoang et al. 1996). Several polymorphic markers (RFLP haplotypes and VNTR and short-tandem-repeat [STR] alleles) have been defined within the gene, adding versatility in the study of the genetic heterogeneity of PKU. The analysis of these intragenic polymorphisms is of particular relevance in population-genetic studies and elucidation of the origins of mutations. During the past few years, there has been a dramatic increase in the number of novel mutations identified, as new molecular approaches for mutation detection are employed. Particularly, the use of denaturing gradient gel electrophoresis (DGGE) and direct sequencing has allowed the characterization of 98% of the mutant alleles in the Danish and Sicilian populations (Guldberg and Güttler 1994). The modification of the method, termed "broad-range" DGGE, adapted to mutation scanning of the PAH gene, allows simultaneous one-step analysis of the 13 PCR-amplified genomic fragments covering the entire coding sequence and splice signals (Guldberg and Güttler 1993). This gene-scanning method is described as the most appropriate and rapid for genotype determination, especially in markedly heterogenous populations such as those in southern Europe.

In previous studies of the Spanish population, 62% of the mutant PAH alleles were identified in 37 PKU patients, by use of a combined approach based on the screening of known mutations and DNA sequencing after single-strand conformation polymorphism (SSCP) (Pérez et al. 1994). A marked difference from the spectrum of mutations in northern Europe was observed. Twenty different mutations were characterized, the

Received August 5, 1996; accepted for publication October 24, 1996.

Address for correspondence and reprints: Dr. Magdalena Ugarte, Centro de Biología Molecular "Severo Ochoa," CSIC-UAM, Facultad de Ciencias, Universidad Autónoma de Madrid, 28049 Madrid, Spain. E-mail: mugarte@mvax.cbm.uam.es

0002-9297/97/6001-0015$02.00

Am. J. Hum. Genet. 60:95–102, 1997

three most frequent ones (IVS10nt-11g→a, I65T, and V388M) accounting only for 30% of the mutant alleles, which confirmed the heterogeneity expected on the basis of previous haplotype studies (Desviat et al. 1993). In order to complete the molecular characterization of the Spanish PKU alleles, we have employed the broad-range DGGE method to localize the remaining mutations prior to sequencing. A systematic analysis of RFLP haplotypes and of VNTR and STR alleles in the PAH gene (Eisensmith and Woo 1992; Goltsov et al. 1992, 1993) associated with the identified mutations also has been conducted. The results highlight the genetic heterogeneity of PKU in the Spanish population and provide a reference for future studies in related populations.

Subjects and Methods

Sixty-three Spanish PKU families representing 129 independent chromosomes were included in this study. In one family, both parents have PKU (Ugarte et al. 1980), representing four independent mutant chromosomes. Another family shows maternal hyperphenylalaninemia in which the father is a heterozygous carrier, adding three independent chromosomes. Siblings with identical genotype were regarded as one patient. Among the patients, two are of Gypsy descent, and one is Moroccan. Most of the patients have been followed up at our laboratory in Madrid. The geographic origin of the patients is spread throughout different regions, so the panel of patients can be considered representative of the Spanish population.

Patients were considered as having PKU when the serum phenylalanine levels at diagnosis were >600 μmol/liter and after exclusion of a defect in tetrahydrobiopterine metabolism. Among the patients, 26% had classical PKU, 16% had moderate PKU, 35% had mild PKU, and 23% corresponded to non-PKU HPA, according to phenylalanine tolerance and plasma phenylalanine levels at diagnosis as described by Güttler et al. (1993).

Genomic DNA from patients and parents was extracted from whole blood by standard procedures (John et al. 1991). In some cases, PCR amplification was performed directly, using dried blood spots as the source of DNA (Pérez et al. 1993). RFLP haplotype analysis was performed by PCR methods and/or by Southern blotting and hybridization with a full-length cDNA probe, kindly provided by Dr. S. L. C. Woo. VNTR alleles in the 3′ end of the gene were analyzed as described by Goltsov et al. (1992). In cases in which the lack of sample precluded the examination of the *Eco*RI and *Eco*RV polymorphisms, the corresponding haplotypes were inferred both from the remaining polymorphisms studied and from the mutation/haplotype association observed in other populations and in ours. To aid in the haplotype determination, certain silent polymorphisms (Q232Q, V245V, and L385L) were examined (Lichter-Konecki et al. 1994).

STR alleles in intron 3 of the PAH gene were amplified by use of the primers described by Goltsov et al. (1993), one of them fluorescently labeled with Cy5. The different allele lengths were resolved on an automated ALF express DNA sequencer using Fragment Manager software (Pharmacia). The numbering of the STR alleles reflects the exact fragment lengths as described by Zschocke et al. (1994).

The 13 genomic fragments covering the entire coding region and splice junctions of the PAH gene were amplified by use of the primers described elsewhere, one of them synthesized with a GC clamp on its 5′ end (Guldberg et al. 1993*b*). The PCR products were analyzed by simultaneous broad-range DGGE (Guldberg and Güttler 1993). All fragments displaying an aberrant migrating band pattern in the DGGE gel were subjected to direct sequencing using the *f*mol sequencing kit from Promega. The sequencing products were analyzed on an automated DNA sequencer (ALF express; Pharmacia). When possible, the identified mutations were confirmed by restriction analysis (either directly or after creation of the restriction site in the amplification) of PCR products amplified again in affected subjects and their parents. The Mendelian inheritance was confirmed, in all cases, by DGGE or restriction-enzyme digestion.

Results

Mutational Spectrum

In an initial study of the molecular basis of PKU in Spain, 62% of the mutant alleles were characterized in 37 patients (Pérez et al. 1994). We now have extended the study to a total of 63 PKU families from different regions of Spain, implementing the use of broad-range DGGE to localize the mutations prior to direct sequencing. Using this approach, we have identified 40 different mutations, which account for close to 90% (115/129) of the total number of mutant alleles. The defined mutations are distributed over all exonic regions of the PAH gene, with the exception of exon 13. Approximately 62% of the mutations are localized in exons 7–12, and, among these, 44% occur in exon 7. The mutational spectrum includes 30 missense mutations, 2 nonsense mutations, 2 deletions, 5 potential splice mutations, and 1 mutation (IVS10nt-11g→a), which has been demonstrated to affect splicing (Dworniczack et al. 1991). The profile of mutation types is similar to the overall worldwide profile (Hoang et al. 1996). Seven mutations—P122Q, D129G, P147S, D151G, A165T, S196fs, and P407S—have not been reported previously. The relative frequencies are shown in table 1. The most frequent mutations are IVS10nt-11g→a (14.7%), I65T (8.5%),

Table 1

Frequencies of PKU Mutations Identified in 129 Independent Spanish Chromosomes with Associated Haplotypes and STR Alleles

Mutation[a]	Frequency (*n*) (%)	Haplotype (*n*)[b]	STR (*n*)[c]
IVS10nt-11g→a	14.7 (19)	6.7 (16)	246 (6), 250 (5), 254 (2), 238 (1), 258 (1), 262 (1)
		34.7 (3)[d]	230 (3)[d]
I65T	8.5 (11)	9.8 (8)	246 (6), 250 (1)
		10.8 (2)	242 (1), 246 (1)
		15.8 (1)	250 (1)
V388M	6.2 (8)	1.7 (6)	242 (1), 238 (2)
		1.8 (2)	238 (1)
R261Q	3.9 (5)	1.8 (5)	238 (2), 246 (1)
E280K	3.9 (5)	1.9 (3)	234 (3)
		1.8 (2)	ND
IVS1nt5	3.1 (4)	4.3 (4)	238 (4)
A403V	3.1 (4)	1.8 (4)	246 (2), 242 (1), 238 (1)
R68S	2.3 (3)	10.8 (1)	242 (1)
		1.8 (2)	246 (1)
R243X	2.3 (3)	4.3 (1)	234 (1)
		1.9 (1)	234 (1)
R243Q	2.3 (3)	27 (1)	ND
		1.8 (2)	246 (1), 238 (1)
P281L	2.3 (3)	1.8 (1)	242 (1)
		1.7 (2)	242 (2)
A309V	2.3 (3)	1.8 (2)	238 (2)
		1.9 (1)	238 (1)
L311P	2.3 (3)	17.9 (1)	234 (1)
		7.9 (2)	234 (2)
S349P	2.3 (3)	1.7 (2)	238 (2)
		1.8 (1)	ND
Y414C	2.3 (3)	4.3 (3)	238 (3)
D415N	2.3 (3)	1.8 (2)	242 (2)
		7.8 (1)	226 (1)
P122Q	1.5 (2)	53.3 (1)	ND
		26.8 (1)	246 (1)
D129G	1.5 (2)	X.9 (2)	226 (1), 238 (1)
R176L	1.5 (2)	4.3 (2)	246 (2)
S196fs	1.5 (2)	X.7 (1)	246 (1)
		X.8 (1)	ND
Y277D	1.5 (2)	5.9 (2)	250 (2)
IVS8nt-7	1.5 (2)	4.3 (2)	242 (2)
A300S	1.5 (2)	1.8 (2)	238 (1), 242 (1)
T380M	1.5 (2)	X.9 (2)	234 (1), 242 (1)
L48S	.8 (1)	4.3 (1)	234 (1)
F39L	.8 (1)	24.3 (1)	242 (1)
P147S	.8 (1)	4.3 (1)	234 (1)
D151G	.8 (1)	1.9 (1)	234 (1)
A165T	.8 (1)	1.8 (1)	242 (1)
Y198fs	.8 (1)	5.9 (1)	250 (1)
G218V	.8 (1)	1.8 (1)	246 (1)
Y204X	.8 (1)	X.12 (1)	230 (1)
P244L	.8 (1)	12.12 (1)	242 (1)
V245A	.8 (1)	3.8 (1)	238 (1)
R252W	.8 (1)	X.3 (1)	242 (1)
IVS7nt1g→a	.8 (1)	1.8 (1)	234 (1)
IVS7nt3g→c	.8 (1)	4.3 (1)	238 (1)
Q304Q	.8 (1)	1.8 (1)	242 (1)
R408Q	.8 (1)	1.9 (1)	242 (1)
P407S	.8 (1)	X.7 (1)	ND
Total no. of alleles identified	89.0 (115)		

NOTE.—Numbers in parentheses, (*n*), are number of alleles. Because of lack of samples, not all haplotypes and STR alleles could be determined.

[a] Novel mutations identified in the present study are underlined; the others have been reported elsewhere by Hoang et al. (1996) and are available from the PAH Mutation Analysis Consortium database, which is accessible via the Internet (at http://www.mcgill.ca/pahdb).

[b] Designations correspond to the listing described by Eissensmith and Woo (1992), which includes the VNTR alleles. An "X" denotes that haplotype could not be assigned unambiguously, because *Eco*RI and *Eco*RV polymorphisms were missing.

[c] Numbering is different from that of Goltsov et al. (1993) and reflects the exact fragment length (Zschocke et al. 1994). ND = not determined, because of lack of parental samples.

[d] Gypsy allele.

and V388M (6.2%). Thirteen mutations are present, with frequencies of 2%–4%, and the remaining mutations are rare, appearing only on one or two mutant chromosomes (table 1).

Using the approach described, we have characterized the two mutant alleles in 53 patients. Of these patients, 78.5% are compound heterozygotes for two different mutations, and 21.5% are homozygous for one mutation, as observed in other European populations. Only two patients, one homozygous for IVS10nt-11g→a and one homozygous for Y414C, were born to consanguineous parents. In 10 patients, only one mutant allele was identified, whereas, in two patients, we observed no abnormal electrophoretic pattern in the 13 exonic fragments analyzed. To discard an inefficiency of the applied broad-range DGGE technique, the analysis was repeated for the 12 patients not completely genotyped, by use of the electrophoretic conditions most suitable for each amplified segment, as described by Guldberg et al. (1993*a*). Heteroduplexes between patient DNA and normal control DNA for all 13 exonic fragments were generated in an attempt to improve the resolution. Finally, the entire exonic and flanking intronic sequences were sequenced in the two patients with no identified mutations and in eight patients with only one characterized allele, confirming the negative results of the DGGE analysis.

Novel Mutations

The seven previously unreported sequence alterations in the PAH gene correspond to six point mutations, causing an amino acid substitution in the mutant protein, and one 22-bp deletion, resulting in a frameshift (table 2). The novel mutation P122Q was found on two chromosomes, on the background of haplotypes 26.8 and 53.3. The D129G mutation was detected on two chromosomes not completely haplotyped. The novel small deletion found in exon 6, S196fs, also was found on two chromosomes not completely haplotyped. The 22-bp deletion corresponds to nucleotides 586–608 in the cDNA (Konecki et al. 1992) and results in a frameshift. The remaining novel mutations were each represented by one chromosome—P147S on haplotype 4.3, D151G on haplotype 1.9, and A165T on haplotype 1.8. The haplotype background of mutation P407S could not be defined, because of lack of parental samples.

Table 2

Novel Mutations Identified in this Study, in Spanish PKU Patients

Nucleotide Change	Codon N/Codon M	Mutation Name	Location
365 C→A	CCA/CAA	P112Q	Exon 4
386 A→G	GAC/GGC	D129G	Exon 4
439 C→T	CCT/TCT	P147S	Exon 4
452 A→G	GAT/GGT	D151G	Exon 5
493 G→A	GCC/ACC	A165T	Exon 5
586del22bp	(586–608)del22bp	S196fs	Exon 6
1219 C→T	CCT/TCT	P407S	Exon 12

Silent Mutations and Polymorphisms

The following known polymorphisms also have been detected in our sample after DGGE analysis: Q232Q, V245V, L385L, IVS3nt-22c→t, and IVS12nt-35c→t. A previously unreported silent mutation 62 bp downstream from exon 1 (IVS1nt+62c→t), which creates a *Bgl*II restriction site, also was identified in 19 mutant alleles, associated with a variety of haplotypes, among them haplotypes 1, 5, 7, and 53. Screening for the Q232Q, V245V, and L385L mutations by restriction-enzyme digestion and/or DGGE analysis showed frequencies of 17% (22/129), 12.4% (16/129), and 5.4% (7/129), respectively. As in other populations, these polymorphisms are found strongly associated with specific haplotypes—Q232Q with haplotypes 3, 4, and 7; V245V with haplotype 4; and L385L with haplotypes 3, 7, and 17. The IVS3nt-22c→t polymorphism was found on haplotypes 4 and 5, and IVS12nt-35c→t was found on haplotype 53.

Associations with Polymorphic Haplotypes and with VNTR and STR Alleles

The RFLP haplotypes and VNTR and STR alleles have been assessed for the mutant chromosomes (table 1). The IVS10nt-11g→a mutation always was found associated with haplotype 6.7, except in two patients, of Gypsy origin, in whom IVS10nt-11g→a is associated with haplotype 34.7 and the 230-bp STR allele. This association is characteristic of all the Spanish PKU Gypsies (Desviat et al., in press). The prevalent STR alleles in the rest of the IVS10nt-11g→a chromosomes are 246 and 250 bp. Mutation I65T was found associated with haplotypes 9 (eight alleles), 10 (two alleles), and 15 (one allele). All the I65T alleles were associated with a VNTR of eight repeats. The prevalent STR allele was 246 bp. There is a high degree of STR variability associated with both mutation IVS10nt-11g→a and mutation I65T. Other mutations found on different haplotype background are R243Q, found on haplotypes 1 and 27; R243X on haplotypes 1 and 4; R68S on haplotypes 1 and 10; L311P on haplotypes 7 and 17; D415N on haplotypes 1 and 7; and the novel mutation P122Q, found on haplotypes 26 and 53. The remaining mutations were found associated with a single RFLP haplotype but, in several cases, with different VNTR and/or STR alleles (table 1).

We have detected 18 different mutations on the background of haplotype 1 and have found 8 mutations on

haplotype 4. Although there are many mutations on haplotype 1, the VNTR and STR alleles are generally different and specific for each mutation, as shown in table 1. Haplotype 5 harbors mutations Y277D and Y198fs, and haplotype 10 harbors mutations I65T and R68S, whereas each of the rest of the haplotypes characterized in our sample harbors only one mutation. The unidentified alleles correspond to haplotype 1 (six alleles), haplotype 27 (two alleles), haplotype 37 (one allele), haplotype 53 (two alleles), and three uncharacterized chromosomes.

Discussion

Detailed information on the mutations causing PAH deficiency within the Spanish population has been obtained. Seven novel mutations have been identified; of these seven, five are localized in exons 4 and 5 of the PAH gene, where, until now, relatively few mutations have been detected (Hoang et al. 1996). Expression analysis of these mutations will reveal their functional consequences on the mutant enzyme and their involvement in the disease phenotype. It can be hypothesized that S196fs is most probably a disease-causing mutation rendering the protein inactive, since it produces a frameshift from nucleotide 586 corresponding to codon 196 of the protein, thus eliminating the normal catalytic core of the enzyme. With this mutation, there are three small deletions (mutations S196fs, L197fs, and Y198fs) occurring in the same region. These deletion events could have occurred because of a slipped-strand mispairing mechanism (Darvasi and Kerem 1995), since short direct repeats are present in the boundaries of the deleted nucleotides (Kleiman et al. 1992).

The mutation profile and relative frequencies reported for the Spanish population confirm the wide discrepancy between northern and southern Europe. In Spain there is a complete absence of the prevalent mutations IVS12nt1 and R408W found in northern and eastern European countries, whereas other mutations, which are frequent in Spain, such as I65T or V388M, are rare in northern Europe (Guldberg et al. 1993*a*) (fig. 1). To achieve a thorough characterization of PKU mutations, their origins, and their functional consequences reflected in the patients' phenotypes, it is necessary to investigate as many populations as possible. With this study, there are, to date, two southern European populations, Spain and Sicily (Guldberg et al. 1993*b*), where the percentage of fully characterized PKU alleles is >90%. Comparing the mutational spectra and relative frequencies in these two populations, we also observe distinct features (fig. 1). Among the prevalent mutations, only IVS10nt-11g→a, the major Mediterranean mutation, is common to both populations. In Sicily, I65T and V388M are virtually absent. I65T is a mutation well represented in Ireland and the British Isles (Tyfield et al. 1993; Zschocke et al. 1995) and present on alleles of Irish or Scottish ("Celtic") origin in Quebec and Australia (Treacy et al. 1993; Ramus et al. 1995). In other European populations, it is present with low frequencies. In Spain, the prevalent STR allele associated with I65T is 246 bp, as has been reported for the Irish and English chromosomes (Ramus et al. 1995; Zschocke et al. 1995), so a common origin is probable. Moreover, it is unlikely that the mutation is recurrent, because a CpG dinucleotide is not involved. The high degree of STR variability associated with the I65T mutation in Spain and in other populations could indicate an ancient origin. The high frequency of I65T in Spain could be due to a number of genetic factors, including genetic drift. The presence of I65T in our population, on the background of haplotypes 10.8 and 15.8, can be explained by a point mutation and/or recombination with a haplotype 9.8 carrying the mutation. The Spanish contribution in Latin American countries in general could account for the presence of this mutation in Chile and Brazil (Pérez et al. 1995, 1996).

V388M is a mutation very frequently found in the Iberian Peninsula, with a higher incidence in Portugal (Leandro et al. 1995), although both the limited number of patients studied in Portugal and the absence of haplotype and STR data preclude the ascertainment of a Portuguese origin. In Spain, V388M is associated exclusively with haplotype 1.7. The study of V388M in Latin America has revealed its presence at a relatively high frequency in Brazil and Chile, with two different origins in Chile, as deduced from the haplotype analysis (Desviat et al. 1995).

The presence of R243Q, R243X, and D415N on two different haplotypes can be explained by a recurrence mechanism, since a CpG dinucleotide is involved (Cooper and Youssoufian 1988). There are other mutations present in our population that segregate with different haplotypes: mutations R68S, L311P, and P122Q (table 1). These haplotypes may have evolved from one another as a result of recombination, gene conversion, or point mutations. The association of a mutation with more than one VNTR and STR allele is quite common (table 1) and already has been observed in other populations, a finding that is attributable to the fact that tandem-repetitive hypervariable DNA sustains a high mutation rate (Jeffreys et al. 1988).

Analysis of the Spanish population has provided, in the past few years, a starting point in the search for PKU mutations in Latin American countries. In this way, screening for common Spanish PKU mutations has resulted in the identification of ~30% of the mutant alleles in countries such as Chile or Brazil (Pérez et al. 1994, 1996). The application of the DGGE gene-scanning method will allow us rapidly to identify the rest of

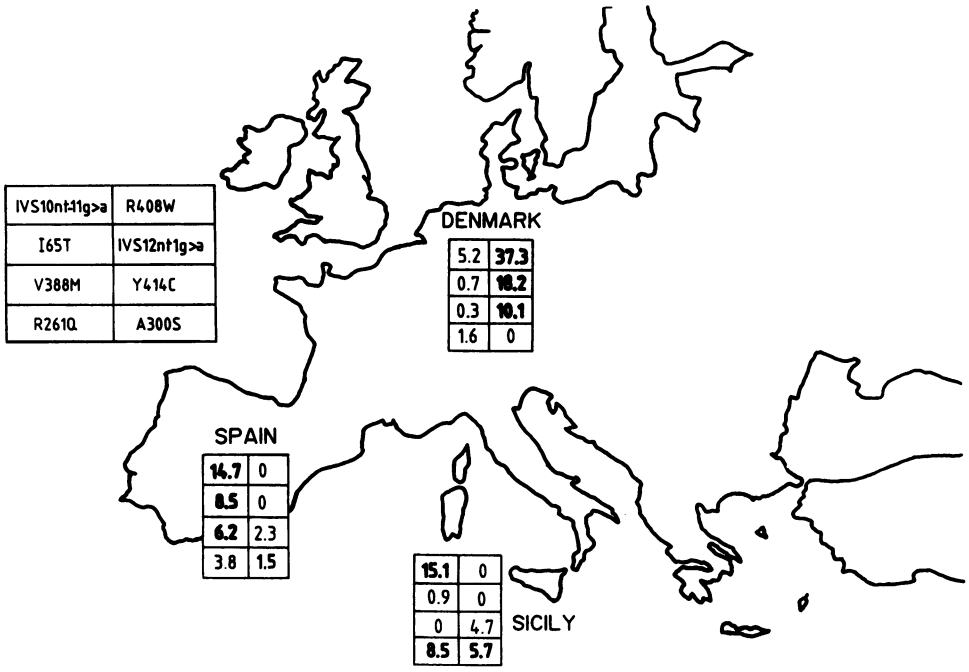

Figure 1 Map of the relative frequency distributions of eight PKU mutations in Spain (present study), Denmark (Guldberg et al. 1993*a*), and Sicily (Guldberg et al. 1993*b*). The three most frequent mutations in each population, denoted in boldface, were chosen for comparison.

the PKU mutations. The examination of the intragenic polymorphisms (haplotypes and VNTR and STR alleles) compared with those detailed in this study will clarify the origins of the mutations in the New World, providing a basis for ascertaining whether they are southern European. In this way, two different origins of V388M in Chile have been proposed (Desviat et al. 1995). Recently, the mutational spectrum of PAH deficiency has been studied in Costa Rica (Santos et al. 1996), revealing the presence of a novel mutation, IVS7nt3, and others found in different populations. On the other hand, mutation IVS10nt-11g→a was not detected in that sample, which led the authors of that study to suggest an absence of Spanish mutations. The absence of IVS10nt-11g→a probably is due to the specific geographic origin of the European ancestors in the group of probands, as occurred in Porto Alegre (Pérez et al. 1996). With the results presented in the present study, it is now clear that the origin of mutations IVS1nt5, L48S, and the novel IVS7nt3, present in patients from Costa Rica, is most probably Spanish. The polymorphisms and VNTR and STR alleles described in Costa Rica correspond to those in Spain that are associated with these mutations—IVS1nt5 and IVS7nt3 on haplotype 4.3 and on the 238-bp STR allele and L48S on haplotype 4.3 and on an STR of 234 bp.

The choice between the different methods for mutation detection may depend on the population in study. In our hands, SSCP analysis and mutation screening allowed us to identify 62% of the mutant alleles (Pérez et al. 1994). The one-step gene-scanning method based on DGGE to localize the mutations prior to sequencing has resulted in a mutation-detection rate of close to 90%, which renders this technique more appropriate for the analysis of populations with marked mutational heterogeneity, such as that in Spain. Identification of the mutations causing PKU in other Spanish patients, referred from different mass-screening and follow-up centers in Spain, now is accomplished easily after DGGE analysis, by comparison of the abnormal migrating pattern obtained and the patterns resulting from the already characterized mutations in our population. We have observed, in some cases, similar band images for two different mutations. These cases can be interpreted by mixing the test sample with a control sample, followed by heteroduplex formation and DGGE. Used in this way, the method is rapid and requires minimum laboratory facilities. Mutations not present in the 126 alleles presently studied will be identified by direct sequencing. The 10% uncharacterized alleles probably will harbor different mutations, since they are present on different haplotypes, among them the highly polymorphic haplotype 1. There are now several PKU alleles in different populations studied, where sequencing of all exonic regions has failed to find a mutation. The fact that the number of unidentified alleles is higher in Spain than in Sicily (Guldberg et al. 1993*b*) and is higher in Sicily than in Denmark (Guldberg et al. 1993*a*) could be due to the

specific genetic heterogeneity present in each population. It has been speculated that the uncharacterized mutations are likely to be situated in the promoter region, at polyadenylation sites, or in the intronic regions of the PAH gene. Both sequencing of such regions and analysis of illegitimate transcripts in our patients not completely genotyped will clarify this point. Big deletions in the genomic DNA can be ruled out in most of our patients, since no abnormal RFLP patterns were observed after Southern blotting and hybridization with a cDNA probe during haplotype analysis.

The data presented in this study will facilitate the establishment of a database of causative southern European PAH mutations and associated intragenic markers, useful for identification and tracing of the migration of mutant alleles during recent human history. The increasing knowledge of the disease genotypes related to the patients' phenotypes will improve diagnosis and permit an earlier implementation of dietary therapy.

Acknowledgments

The authors wish to thank A. Sánchez for her excellent technical assistance and the following physicians for sending samples: Dr. M. Martínez-Pardo (Madrid), Dr. A. Baldellou (Zaragoza), Dr. P. Sanjurjo (Bilbao), Dr. J. Arena (San Sebastian), and Dr. M. A. Vilaseca (Barcelona). The institutional grant of 'Fundación Ramón Areces' to the Centro de Biología Molecular "Severo Ochoa" is gratefully acknowledged. This work was supported by grant SAF-93-0076 from the Comisión Interministerial de Ciencia y Tecnología.

References

Cooper DN, Youssoufian H (1988) The CpG dinucleotide and human genetic disease. Hum Genet 78:151–155

Darvasi A, Kerem B (1995) Deletion and insertion mutations in short tandem repeats in the coding regions of human genes. Eur J Hum Genet 3:14–20

Desviat LR, Pérez B, De Lucca M, Cornejo V, Schmidt B, Ugarte M (1995) Evidence in Latin America of recurrence of V388M, a phenylketonuria mutation with high in vitro residual activity. Am J Hum Genet 57:337–342

Desviat LR, Pérez B, Ugarte M (1993) Phenylketonuria in Spain: RFLP haplotypes and linked mutations. Hum Genet 92:254–258

———. Phenylketonuria in Spanish Gypsies: prevalence of the IVS10nt546 mutation on haplotype 34. Hum Mutat (in press)

Dworniczack B, Aulehla-Scholz C, Kalaydjieva L, Bartholomé K, Grudda K, Horst J (1991) Aberrant splicing of phenylalanine hydroxylase mRNA: the major cause of phenylketonuria in parts of southern Europe. Genomics 11:242–246

Eisensmith RC, Woo SLC (1992) Updated listing of haplotypes at the human phenylalanine hydroxylase (PAH) locus. Am J Hum Genet 51:1445–1448

Goltsov AA, Eisensmith RC, Konecki DS, Lichter-Konecki U, Woo SLC (1992) Associations between mutations and a VNTR in the human phenylalanine hydroxylase gene. Am J Hum Genet 51:627–636

Goltsov AA, Eisensmith RC, Naughton ER, Jin L, Chakraborty R, Woo SLC (1993) A single polymorphic STR system in the human phenylalanine hydroxylase gene permits rapid prenatal diagnosis and carrier screening for phenylketonuria. Hum Mol Genet 2:577–581

Guldberg P, Güttler F (1993) 'Broad-range' DGGE for single-step mutation scanning of entire genes: application to human phenylalanine hydroxylase gene. Nucleic Acids Res 22: 880–881

——— (1994) Mutation screening versus gene scanning for genotyping phenylketonuria patients. J Inherit Metab Dis 17:359–361

Guldberg P, Henriksen KF, Güttler F (1993*a*) Molecular analysis of phenylketonuria in Denmark: 99% of the mutations detected by denaturing gradient gel electrophoresis. Genomics 17:141–146

Guldberg P, Romano V, Ceratto N, Bosco P, Ciuna M, Indelicato A, Mollica F, et al (1993*b*) Mutational spectrum of phenylalanine hydroxylase deficiency in Sicily: implications for diagnosis of hyperphenylalaninemia in southern Europe. Hum Mol Genet 2:1703–1707

Güttler F, Guldberg P, Henriksen KF, Mikkelsen I, Olsen B, Lou H (1993) Molecular basis for the phenotypical diversity of phenylketonuria and related hyperphenylalaninemias. J Inherit Metab Dis 16:602–604

Hoang L, Byck S, Prevost L, Scriver CR (1996) PAH Mutation Analysis Consortium database: a database for disease-producing and other allelic variation at the human PAH locus. Nucleic Acids Res 24:127–131

Jeffreys AJ, Royle NJ, Wislon V, Wong Z (1988) Spontaneous mutation rates to new length alleles at tandem-repetitive hypervariable loci in human DNA. Nature 332:278–281

John SWM, Weitzner G, Rozen R, Sciver CR (1991) A rapid procedure for extracting genomic DNA from leukocytes. Nucleic Acids Res 19:408

Kleiman S, Schwartz G, Woo SLC, Shiloh Y (1992) A 22-bp deletion in the phenylalanine hydroxylase gene causing phenylketonuria in an Arab family. Hum Mutat 1:344–346

Konecki DS, Wang Y, Trefz FK, Lichter-Konecki U, Woo SLC (1992) Structural characterization of the 5′ regions of the human phenylalanine hydroxylase gene. Biochemistry 31: 8363–8368

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

Leandro P, Rivera I, Ribeiro V, Tavares de Almeida I, da Silveira C, Lechner MC (1995) Mutation analysis of phenylketonuria in south and central Portugal: prevalence of V388M mutation. Hum Mutat 6:192–194

Lichter-Konecki, Schlotter M, Konecki DS (1994) DNA sequence polymorphisms in exonic and intronic regions of the human phenylalanine hydroxylase gene aid in the identification of alleles. Hum Genet 94:307–310

Pérez B, Desviat LR, De Lucca M, Cornejo V, Ugarte M (1995) Mutations and polymorphisms in the phenylalanine hydroxylase gene in Chile. Am J Hum Genet Suppl 57:A170

Pérez B, Desviat LR, De Lucca M, Schmidt B, Loghin-Grosso, Giugliani R, Pires RF, et al (1996) Mutation analysis of phenylketonuria in south Brazil. Hum Mutat 8:262–264

Pérez B, Desviat LR, De Lucca M, Ugarte M (1994) Spectrum and origin of phenylketonuria mutations in Spain. Acta Pediatr Suppl 407:34–36

Pérez B, Desviat LR, Díe M, Cornejo V, Chamoles NA, Nicolini H, Ugarte M (1993) Presence of the Mediterranean PKU mutation IVS10 in Latin America. Hum Mol Genet 2:1289–1290

Ramus SJ, Treacy EP, Cotton RGH (1995) Characterization of phenylalanine hydroxylase gene in untreated phenylketonuric patients from Victoria, Australia: origin of alleles and haplotypes. Am J Hum Genet 56:1034–1041

Santos M, Kuzmin AI, Eisensmith RC, Goltsov AA, Woo SLC, Barrantes R, de Céspedes C (1996) Phenylketonuria in Costa Rica: preliminary spectrum of PAH mutations and their associations with highly polymorphic haplotypes. Hum Hered 46:128–131

Scriver CR, Kaufman S, Eisensmith RC, Woo SCL (1995) The hyperphenylalaninemias. In: Scriver CR, Beaudet AL, Sly WS, Valle D (eds) The metabolic and molecular bases of inherited disease, 7th ed. McGraw-Hill, New York, pp 1015–1075

Treacy E, Byck S, Clow C, Scriver CR (1993) "Celtic" PKU chromosomes found? evidence in two regions of Quebec province. Eur J Hum Genet 1:220–228

Tyfield LA, Osborn MJ, King SK, Jones MM, Holton JB (1993) Molecular basis of phenylketonuria in a English population. Dev Brain Dysfunct 6:60–67

Ugarte M, Maties M, Ugarte JL (1980) The offspring of a phenylketonuric couple. J Ment Defic Res 24:119–127

Zschocke J, Graham CA, Carson DJ, Nevin NC (1995) Phenylketonuria mutation analysis in Northern Ireland: a rapid stepwise approach. Am J Hum Genet 57:1311–1317

Zschocke J, Graham CA, McKnight JJ, Nevin NC (1994) The STR system in the human phenylalanine hydroxylase gene: true fragment length obtained with fluorescent labelled PCR primers. Acta Pediatr Suppl 407:41–42