Am. J. Hum. Genet. 51:1355–1365, 1992

Multiple Origins for Phenylketonuria in Europe

R. C. Eisensmith,* Y. Okano,* M. Dasovich,* T. Wang,* F. Güttler,‡ H. Lou,‡ P. Guldberg,‡ U. Lichter-Konecki,§ D. S. Konecki,§ E. Svensson,‖ L. Hagenfeldt,‖ F. Rey,# A. Munnich,# S. Lyonnet,# F. Cockburn,** J. M. Connor,** M. E. Pembrey,†† I. Smith,†† R. Gitzelmann,‡‡ B. Steinmann,‡‡ J. Apold,§§ H. G. Eiken,§§ M. Giovannini,‖‖ E. Riva,‖‖ R. Longhi,‖‖ C. Romano,## R. Cerone,## E. R. Naughten,*** C. Mullins,*** S. Cahalane,*** I. Özalp,††† G. Fekete,‡‡‡ D. Schuler,‡‡‡ G. Y. Berencsi,§§§ I. Nász,§§§ R. Brdicka,‖‖‖ J. Kamaryt,### A. Pijackova,### B. Cabalska,**** K. Boszkowa,**** E. Schwartz,†††† V. N. Kalinin,‡‡‡‡ L. Jin,† R. Chakraborty,† and S. L. C. Woo*

*Howard Hughes Medical Institute, Department of Cell Biology and Institute for Molecular Genetics, Baylor College of Medicine, and †Genetics Center, University of Texas Graduate School of Biomedical Sciences, Houston; ‡The John F. Kennedy Institute, Glostrup, Denmark; §Universitäts-Kinderklinik, Heidelberg; ‖Department of Clinical Chemistry, Huddinge University Hospital, Huddinge, Sweden; #Unité de Recherches sur les Handicaps Génétique de l'Enfant, Hôpital des Enfants-Malades, Paris; **Department of Medical Genetics, Duncan Guthrie Institute of Medical Genetics, University of Glasgow, Glasgow; ††Institute of Child Health, University of London, London; ‡‡Universitäts-Kinderklinik, Zürich; §§Avdeling for Medisinsk Genetikk, Universitet i Bergen, Bergen; ‖‖Clinica Pediatrica V^a, Università Degli Studi Di Milano, Milan; ##Clinica Pediatrica, Università Di Genova, Centro Regionale di Riferimentao Per Le Malattie Endocrine E Metaboliche Dell'Età Pediatrica, Genoa; ***Metabolic Unit, The Children's Hospital, Dublin; †††Department of Metabolism, Institute of Child Health, Hacettepe University, Ankara; ‡‡‡Department of Paediatrics No. 2 and §§§Department of Microbiology, Semmelweis University Medical School, Budapest; ‖‖‖Institute of Hematology and Blood Transfusion, Prague: ###Research Institute of Pediatrics, Brno, Czech and Slovak Federal Republic; ****Instytut Matki i Dziecka, L'Institut National de la Mere et de L'enfat, Warsaw; ††††Center for Molecular Diagnostics and Laboratory of Human Molecular Genetics, St. Petersburg Institute of Nuclear Physics, Gatchina, Russian Republic; and ‡‡‡‡Laboratory of Molecular Genetics, National Research Centre of Medical Genetics, Moscow

Summary

Phenylketonuria (PKU), a disorder of amino acid metabolism prevalent among Caucasians and other ethnic groups, is caused primarily by a deficiency of the hepatic enzyme phenylalanine hydroxylase (PAH). PKU is a highly heterogeneous disorder, with more than 60 molecular lesions identified in the PAH gene. The haplotype associations, relative frequencies, and distributions of five prevalent PAH mutations (R158Q, R261Q, IVS10nt546, R408W, and IVS12nt1) were established in a comprehensive European sample population and subsequently were examined to determine the potential roles of several genetic mechanisms in explaining the present distribution of the major PKU alleles. Each of these five mutations was strongly associated with only one of the more than 70 chromosomal haplotypes defined by eight RFLPs in or near the PAH gene. These findings suggest that each of these mutations arose through a single founding event that occurred within time periods ranging from several hundred to several thousand years ago. From the significant differences observed in the relative frequencies and distributions of these five alleles throughout Europe, four of these putative founding events could be localized to specific ethnic subgroups. Together, these data suggest that there were multiple, geographically and ethnically distinct origins for PKU within the European population.

Received May 19, 1992; final revision received September 11, 1992.

Address for correspondence and reprints: R. C. Eisensmith, Ph.D., Department of Cell Biology, Baylor College of Medicine, Houston, Texas 77030.

0002-9297/92/5106-0018$02.00

Introduction

Phenylketonuria (PKU) is an autosomal recessive genetic disorder caused primarily by a deficiency of hepatic phenylalanine hydroxylase (PAH). PAH converts L-phenylalanine to L-tyrosine, and severe PAH

deficiency results in an accumulation of L-phenylalanine in the serum. The primary symptom in individuals with PKU is severe, irreversible mental retardation. The frequency of PKU ranges from approximately 1 in 2,600 in Turkey (Özalp et al. 1986) to approximately 1 in 120,000 in Japan (Aoki and Wada 1988). Overall, the frequency among Caucasians is approximately 1 in 10,000 (Bickel et al. 1981), corresponding to a carrier frequency of about 1 in 50 (for review, see Scriver et al. 1989). Extensive RFLPs have been identified at the human PAH locus (Woo et al. 1983; Lidsky et al. 1985). The differences in the relative frequencies and distributions of various RFLP haplotypes among normal and mutant PAH genes in different European populations provided the foundation for molecular genetic analysis of PKU among Caucasians. The stong associations between some haplotypes and mutant chromosomes suggested that these haplotypes contained a single, predominant PAH mutation, similar to previous observations at the β-globin locus (for review, see Orkin and Kazazian 1984).

Direct molecular analysis of mutant PAH chromosomes has since identified the IVS12nt1 splicing mutation on mutant haplotype 3 chromosomes (DiLella et al. 1986), the R408W missense mutation on mutant haplotype 2 chromosomes (DiLella et al. 1987), and the IVS10nt546 splicing mutation on chromosomes of mutant haplotype 6 (Dasovich et al. 1991; Kalaydjieva et al. 1991) and some other haplotypes (Dworniczak et al. 1991*a;* Kalaydjieva et al. 1991). Similar studies on chromosomes of haplotypes well represented among both normal and mutant chromosomes have identified at least 14 putative PAH mutations associated with mutant haplotype 1 chromosomes and at least 15 putative mutations associated with mutant haplotype 4 chromosomes (for reviews, see John et al. 1990; Eisensmith and Woo 1991, 1992*a;* Konecki and Lichter-Konecki 1991; Scriver et al., in press). The most common mutation yet observed on haplotype 1 chromosomes in European populations is the R261Q mutation (Abadie et al. 1989; Okano et al. 1990), while the most common mutation yet observed on haplotype 4 chromosomes is the R158Q mutation (Dworniczak et al. 1989; Okano et al. 1990).

Since the discovery of PKU more than 50 years ago (Fölling 1934), one interesting question has been how this autosomal recessive disorder attained its present frequency in human populations. The nearly absolute associations initially observed between certain PAH mutations and specific RFLP haplotypes in discrete European populations suggested that the present distribution of mutant PAH chromosomes in Europe may be due to founder effect and drift. To test this hypothesis, we determined the relative frequencies and distributions of the major PAH mutations in a large number of European sample populations. These data were then used to examine the potential roles of several genetic mechanisms in explaining the present distribution of the major PKU alleles.

Material and Methods

Sample Collection

Venous blood samples from nuclear PKU families were collected at 19 different clinics in 10 different European populations. Although nearly every hyperphenylalaninemic patient detected in the nationwide newborn screening program was included in the Swedish sample population, only patients with mild or classical PKU collected by one or more regional centers were included in the other sample populations. In addition, the number of samples collected within each of these countries was generally not proportional to the the incidence of PKU, or to the size of these populations, relative to the total European population. Thus, only relative frequencies could be established, with some possibility for sampling errors.

Determination of Normal and Mutant RFLP Haplotypes at the PAH Locus

RFLP analysis of genomic DNA isolated from venous blood samples was performed by digestion with seven specific restriction endonucleases, followed by Southern blot analysis using the full-length human PAH cDNA clone (phPAH 247) as a hybridization probe. Haplotypes of individual chromosomes were determined from the RFLPs as described elsewhere (Chakraborty et al. 1987). Haplotype data for normal chromosomes were derived from studies of those normal chromosomes present in PKU families, rather than in unaffected families, raising the possibility of some bias. However, previous haplotype analyses of the PAH and other loci have demonstrated that frequencies derived from this approach are unbiased estimates of their population values (Chakraborty 1986; Chakraborty et al. 1987). Associations between mutations and haplotypes were quantified by both D/D_{max} (Lewontin 1964) and D_s (Hill and Robertson 1968). The significance of the associations between normal or mutant chromosomes and RFLP haplotypes, and between PAH mutations and RFLP haplotypes, were determined by Fisher's exact test.

Detection of Known PAH Mutations by Allele-specific Oligonucleotide (ASO) Hybridization

The frequencies of five prevalent PAH mutations were determined in each European sample population by PCR amplification of the exonic regions of the PAH gene, followed by slot-blot hybridization using ASO probes according to conditions reported elsewhere (DiLella et al. 1988). All available samples from European populations were included in these experiments, regardless of whether the haplotypes were known or unknown. Thus, mutation frequency data were available on samples from English, Irish, Polish, Russian, and other populations that were not included in the haplotype studies.

Results

Relative Frequencies and Distributions of PAH RFLP Haplotypes

Of the 71 PAH RFLP haplotypes recorded to date (Woo 1988; Eisensmith and Woo 1992*b*), only 7 (haplotypes 1–7) accounted for over 80% of all mutant PAH chromosomes in most European populations (table 1). Haplotypes 3 and 6 were strongly associated with mutant chromosomes in several northern and southern European populations, respectively, while haplotype 2 was strongly associated with mutant chromosomes in many European populations. Haplotypes 5 and 7 were more often present on normal chromosomes in several European populations, especially in Germany, but haplotype 7 was associated with mutant chromosomes in Norway. Although haplotype 4 was more often found on normal chromosomes in some European populations, this haplotype and haplotype 1 were common among both normal and mutant chromosomes in most European populations.

Haplotype Associations, Relative Frequencies, and Distributions of the Major PAH Mutations

Although over 60 mutations have been identified in the human PAH gene, only about 10 of these are present on a significant proportion of mutant chromosomes of haplotypes 1–7 in Caucasian populations (for review, see Eisensmith and Woo 1992*a*). The locations of five of these mutations (R158Q, R261Q, IVS10nt546, R408W, and IVS12nt1) are shown schematically in figure 1. Significant associations, as quantified by D/D_{max} (Lewontin 1964) and D_s (Hill and Robertson 1968), existed between these five PAH mutations and specific RFLP haplotypes in most European populations (table 2). The relative frequencies and distributions of these mutations in 14 European sample populations are presented in table 3.

Several different types of mutation/haplotype association were evident in this study. In one type of association, a specific mutation was linked to a single haplotype, but not all chromosomes of this haplotype contained a specific mutation. This type of mutation/haplotype association is termed "exclusive," and the exclusive nature of this association is reflected in a perfect D/D_{max} value (D/D_{max} value = 100; table 2). In the second type of association, a specific mutation was again linked to a single haplotype, but in this case all chromosomes of this haplotype contained this mutation. This type of mutation/haplotype association is not only exclusive but is also termed "inclusive." The inclusive nature of this association is evident from a perfect D_s value (D_s = 100; table 2).

The IVS12nt1 mutation.—This mutation was present on about 95% of all mutant haplotype 3 chromosomes and on less than 0.5% of all non–haplotype 3 chromosomes (table 2). Thus, chromosomes containing the IVS12nt1 mutation included nearly all mutant haplotype 3 chromosomes, and vice versa. Moreover, this association was strongly exclusive of other mutations and haplotypes, in that the IVS12nt1 mutation was almost never found on chromosomes of other haplotypes; nor are other mutations found on the vast majority of mutant haplotype 3 chromosomes. The frequency gradient, shown in figure 2 (values in upper left quadrants), suggested that this mutation first occurred on a normal haplotype 3 chromosome in a Danish founding population and was subsequently spread into neighboring populations. The strong association observed between this mutation and haplotype 3 suggested that this founding event occurred within the past several hundred to few thousand years. It is unlikely that this frequency gradient was an artifact of sampling, since both newborn screening and sample collection are quite extensive within the Scandinavian countries where this mutation was most prevalent and where this gradient was already evident.

The R408W mutation.—The association between the R408W mutation and haplotype 2 was also strongly inclusive, since this mutation was present on over 90% of all haplotype 2 chromosomes (table 2). This association was, however, less exclusive of other haplotypes, since the R408W mutation was also observed on a number of haplotype 1 chromosomes, as first described by John et al. (1990). The R408W mutation

Table 1

Relative Frequencies and Distributions of PAH Haplotypes in 14 European Sample Populations

Region and Country	Haplotype 1		Haplotype 2		Haplotype 3		Haplotype 4		Haplotype 5		Haplotype 6		Haplotype 7		Haplotypes 1–7		No. of Alleles	
	Normal	Mutant	Normal	Mutant	Normal	Mutant	Normal	Mutant	Normal	Mutant	Normal	Mutant	Normal	Mutant	Normal	Mutant	Normal	Mutant
Northern Europe:																		
Denmark[a]	35[b]	18	4	20[b]	3	38[c]	32[b]	13	10[b]	0	0	3	10[b]	2	94	94	66	66
Norway[d]	28	21	7	11	4	17[e]	27[e]	9	14	6	2	0	8	20[b]	92	84	85	86
Sweden[f]	31[f]	15	8	23[c]	2	14[c]	19	27	8	6	12[c]	1	7	4	76	90[e]	132	136
Western Europe:																		
France[g]	26	31	6	18[b]	3	9	15	8	8[b]	0	2	3	12	4	72	73	68	74
Germany[h]	24	22	8	32[c]	3	13[c]	20	19	12[c]	2	2	2	14[c]	2	83	92[e]	246	246
Scotland[i]	32	30	3	9	6	18	13	6	10	3	3	0	10	0	77	66	31	33
Switzerland[i]	42	50	6	11	0	5	25	18	11	0	0	5	11	0	95	89	36	38
Eastern Europe:																		
Bulgaria[j]	24	16	2	38[c]	2	0	15[b]	10	2	2	2	22[e]	10[b]	0	57	88[e]	42	50
Czechoslovakia[k]	14	0	33	68[b]	5	0	14	23	5	0	0	0	0	0	71	91	21	22
Hungary[k]	32	13	16	55[e]	8	5	16	10	0	2	0	0	5	0	77	85	38	40
Poland[l]	27	9	9	57[c]	0	2	11	11	9	2	0	5	7	5	63	91[e]	44	44
Southern Europe:																		
Greece[m]	30	47	10	18	0	0	20	29	NA	NA	NA	NA	NA	NA	⩾60	⩾94[b]	10	17
Italy[n]	29	40	0	6	3	3	24[b]	9	9	4	2	18[e]	5	0	72	80	63	68
Turkey[o]	14	25	7	1	2	1	31[b]	17	10	4	2	36[c]	11[b]	2	78	86	84	91

[a] Chakraborty et al. (1987).

[b] Significant disequilibrium between normal or mutant chromosome and haplotype ($P < .05$, by Fisher's exact test).

[c] Significant disequilibrium between normal or mutant chromosome and haplotype ($P < .001$, by Fisher's exact test).

[d] Apold et al. (1990).

[e] Significant disequilibrium between normal or mutant chromosome and haplotype ($P < .005$, by Fisher's exact test).

[f] Svensson et al. (1991).

[g] Rey et al. (1988).

[h] Aulehla-Scholz et al. (1988), Herrmann et al. (1988), Lichter-Konecki et al. (1988), and Riess et al. (1988).

[i] Sullivan et al. (1989).

[j] Kalaydjieva et al. (1990).

[k] Daiger et al. (1989*a*).

[l] Zygulska et al. (1991).

[m] Hofman et al. (1989). NA = not available; these data were combined into a single category representing haplotypes 5–12, rather than reported separately.

[n] Dianzani et al. (1990).

[o] Lichter-Konecki et al. (1989) and Stuhrmann et al. (1989).

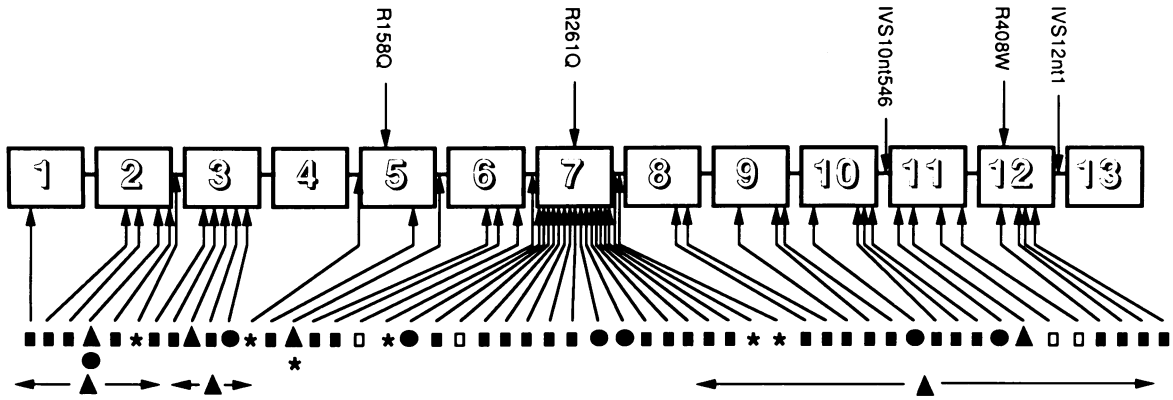

Figure 1 Mutations in the human PAH gene. The locations of five PAH mutations common in some European populations are shown above the schematic representation of the exonic structure of the gene. The relative locations of other known PAH mutations are shown at the bottom of this diagram. The blackened squares (■) denote the presence of missense mutations, the blackened circles (●) denote the presence of nonsense mutations, the asterisks (*) denote the presence of splicing mutations, the blackened triangles (▲) denote the presence of deletions, and the unblackened squares (□) denote the presence of polymorphisms.

was most frequent in eastern European countries, where it accounted for about 50% or more of all mutant alleles (table 3 and fig. 2, values in upper right quadrants). Since haplotype 2 was most common among chromosomes in Czechoslovakia, it is more likely that the R408W mutation initially occurred on a haplotype 2 chromosome within this population, although the absence of haplotype and frequency data from the more eastern regions of the Russian and other republics of the former Soviet Union precluded a truly precise localization of a putative founding population. The absence of this mutation from haplotype 2 chromosomes in Chinese or Japanese populations (T. Wang, unpublished observation) suggested that the founding event that introduced the R408W mutation onto haplotype 2 chromosomes was unique to caucasoid peoples. The strong association still present between this mutation and haplotype 2 suggested that this founding event also occurred within the past few millennia.

The IVS10nt546 mutation.—The association between the IVS10nt546 mutation and haplotype 6 was less exclusive of other mutations, since this mutation was present on only 85% of haplotype 6 chromosomes in Europe, but was strongly exclusive of other haplotypes, since this mutation was present on less than 0.5% of all non–haplotype 6 chromosomes in these sample populations (table 2). Similar to the previous mutations examined, the strong mutation/haplotype association suggested that the IVS10nt546 mutation also occurred fairly recently, within the past few thousand years. IVS10nt546 was the most common mutation yet reported in southern European populations (table 3). This mutation has also been observed in the Spanish population, where it accounts for 11% of all mutant alleles (Perez et al. 1992), and on a majority of the mutant haplotype 6 chromosomes in the Bulgarian population (Kalaydjieva et al. 1991). The distribution of this mutation (fig. 2, values in lower left quadrants) was certainly suggestive of a Turkish origin, with a subsequent spread throughout the Mediterranean basin. However, a more recent examination of the distribution of this mutation within the Italian population has shown that this allele is present primarily in regions that had been settled by Italian peoples prior to 1000 B.C. and not in regions settled by Turks or other Middle Eastern groups (Romano et al., in press), suggesting an Italian founding population for this mutation. The frequency of this allele in the Turkish population would then be a reflection of the high degree of consanguinity which has increased the incidence of PKU in this population (Özalp et al. 1986).

The R261Q mutation.—The association between the R261Q mutation and haplotype 1 was strongly exclusive, since this mutation was not present on chromosomes of other haplotypes in these sample populations, but was not inclusive of all mutant haplotype 1 chromosomes. This lack of inclusivity suggested that the event creating the R261Q mutation occurred well after the events mediating the origin of haplotype 1. Since this mutation was relatively frequent in both Switzerland and Turkey but rare in most other populations (table 3 and fig. 2, values in lower right quadrants), it could either have occurred in a single found-

Table 2

Associations between PAH Mutations and RFLP Haplotypes

Region and Country	R261Q				R408W				IVS12nt1				R158Q				IVS10nt546			
	Haplotype 1	Non-Haplotype 1	D/D_{max}	D_s	Haplotype 2	Non-Haplotype 2	D/D_{max}	D_s	Haplotype 3	Non-Haplotype 3	D/D_{max}	D_s	Haplotype 4	Non-Haplotype 4	D/D_{max}	D_s	Haplotype 6	Non-Haplotype 6	D/D_{max}	D_s
Northern Europe:																				
Denmark	3/30	0/100	100	28[a]	20/21	1/109	94	94[b]	40/40	0/90	100	100[b]	4/19	0/111	100	43[b]	6/8	0/122	100	86[b]
Norway	7/17	0/77	100	22[b]	7/9	8/85	74	55[b]	16/18	2/76	86	86[b]	2/9	0/85	100	45[c]	0/0	0/94	...	...
Sweden	2/20	0/116	100	29[a]	30/31	2/105	96	94[b]	18/19	0/117	100	97[b]	2/37	0/99	100	20	0/2	0/134	...	...
Western Europe:																				
France	10/25	0/45	100	55[b]	8/9	1/61	87	87[b]	9/9	0/61	100	100[b]	ND	ND	ND	ND	ND	ND	ND	ND
Germany	12/40	0/45	100	43[b]	52/52	0/49	100	100[b]	32/32	0/34	100	100[b]	12/47	0/30	100	34[d]	5/5	2/109	100	84[b]
Switzerland	13/18	0/18	100	75[b]	3/4	0/32	100	85[b]	2/2	0/34	100	100[d]	2/6	0/30	100	54[a]	1/2	0/34	100	70
Eastern Europe:																				
Czechoslovakia	0/0	0/20	...	...	12/14	0/6	100	80[b]	0/0	0/20	...	...	2/5	0/15	100	58	0/0	0/20	...	...
Hungary	1/5	0/36	100	42	21/22	1/19	90	90[b]	3/3	0/38	100	100[b]	4/5	0/36	100	88[b]	0/0	0/41	...	...
Southern Europe:																				
Italy	2/20	0/28	100	25	1/4	0/44	100	48	1/4	0/44	100	48	0/1	0/47	...	...	2/5	0/43	100	61[c]
Turkey	8/22	0/12	100	41[a]	1/1	0/4	100	100	0/1	0/3	...	...	4/19	0/5	100	23	31/31	0/28	100	100[b]

NOTE.—Only data from countries where haplotyped samples were screened for these PAH mutations are reported. Fractions are number of positive alleles divided by total number of alleles examined. ND = not determined.

[a] Significant disequilibrium between mutation and haplotype ($P < .05$ by Fisher's exact test).

[b] Significant disequilibrium between mutation and haplotype ($P < .001$ by Fisher's exact test).

[c] Significant disequilibrium between mutation and haplotype ($P < .01$ by Fisher's exact test).

[d] Significant disequilibrium between mutation and haplotpye ($P < .005$ by Fisher's exact test).

Table 3

Relative Frequencies and Distributions of PAH Mutations in Europe

Region and Country	R158Q[a]	R261Q	IVS10nt546	R408W	IVS12nt1
Northern Europe:					
Denmark	4/150 (2.7)	4/150 (2.7)	6/150 (4.0)	32/150 (21.3)	67/150 (44.7)
Norway	2/94 (2.1)	7/94 (7.4)	0/94 (0)	15/94 (16.0)	18/94 (19.0)
Sweden	3/178 (1.7)	3/178 (1.7)	0/178 (0)	39/178 (21.9)	28/178 (15.7)
Western Europe:					
England	2/38 (5.3)	2/38 (5.3)	0/38 (0)	4/38 (10.5)	6/38 (15.8)
France	ND	10/127 (7.8)	7/88 (8.0)	9/127 (7.0)	16/127 (12.5)
Germany	12/202 (5.9)	12/202 (5.9)	5/202 (2.5)	52/202 (25.7)	32/202 (15.8)
Ireland	0/36 (0)	0/36 (0)	5/36 (13.9)	13/36 (36.1)	1/36 (2.8)
Switzerland	2/50 (4.0)	16/50 (32.0)	2/50 (4.0)	3/50 (6.0)	2/50 (4.0)
Eastern Europe:					
Czechoslovakia	2/36 (5.6)	1/36 (2.8)	0/34 (0.0)	22/36 (61.1)	0/36 (0)
Hungary	5/70 (7.1)	1/70 (1.4)	2/70 (2.9)	34/70 (48.6)	3/70 (4.3)
Poland	ND	0/26 (0)	1/26 (3.8)	17/26 (65.4)	0/26 (0)
Russia	8/220 (3.6)	2/156 (1.3)	4/330 (1.2)	133/218 (61.0)	3/156 (1.9)
Southern Europe:					
Italy	0/72 (0)	2/72 (2.8)	9/72 (12.5)	1/72 (1.4)	2/72 (2.8)
Turkey	4/83 (4.8)	8/83 (9.6)	31/79 (39.2)	1/83 (1.2)	0/83 (0)

NOTE.—Fractions are number of positive alleles divided by total number of alleles examined.
[a] ND = not determined.

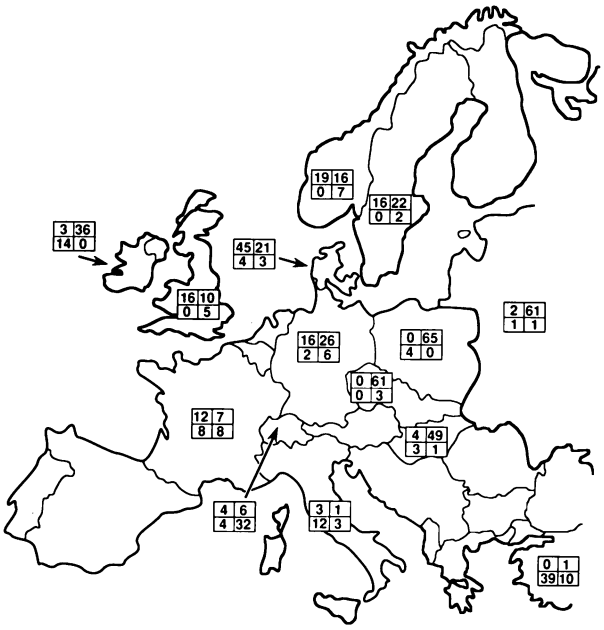

Figure 2 Map of the relative frequency distributions of the four PAH mutations common to some European populations. The numbers in the four quadrants of each box correspond to the percentages of PKU chromosomes containing the IVS12nt1 (upper left), R408W (upper right), IVS10nt546 (lower left), and R261Q (lower right) mutations.

ing population—with its frequency increased in Switzerland and, to a lesser extent, in Turkey through genetic drift—or it could have occurred independently in these two populations. The strong association of this mutation with haplotype 1, especially in Turkey, where haplotype 1 was relatively rare among normal alleles, suggested that recurrence is unlikely. This conclusion is further supported by the more recent finding of a strong association between this mutation and a single VNTR allele present on the polymorphic haplotype 1 background in these two populations (Goltsov et al. 1992).

The R158Q mutation.—The association of the R158Q mutation with haplotype 4 was very similar to that seen between the R261Q mutation and haplotype 1, in that it was absolutely exclusive ($D/D_{max} = 100\%$) but not inclusive (only about 20% of haplotype 4 chromosomes contained the R158Q mutation). Again, since this mutation also involves a CpG dinucleotide, it is unlikely that such a strong mutation/haplotype association could be the result of recurrence of this mutation in different European populations; rather, it reflects the occurrence of a single, relatively recent founding event. In contrast to the other mutations examined in this study, there was no clear spatial pat-

tern that provides any insight into the origin of this mutation (table 3). The effects of genetic drift on the low frequency of this mutation have obliterated all evidence concerning a possible location for a putative founding population.

Discussion

There are five principal means whereby autosomal recessive disorders may achieve high frequencies in humans: (1) founder effect and drift, (2) heterozygote selection, (3) elevated mutation rate, (4) reproductive compensation, and (5) the involvement of multiple loci that confer the same disease phenotype. Multiple loci can be discounted in PKU, as defects in the PAH gene account for more than 95% of all cases. There is also no evidence for reproductive compensation, although possible effects of this mechanism during earlier time periods cannot be completely dismissed. There is evidence that some PKU genotypes are the result of recurrent mutation (John et al. 1990; Okano et al. 1990; Tsai et al. 1990; Dworniczak et al. 1991*b*). In each case, these mutations involve CpG dinucleotides that are apparent hot spots for mutations in the mammalian genome (Cooper and Youssoufian 1988) and the PAH gene (Abadie et al. 1989). However, elevated mutation rate can be discounted as the principal cause of PKU, since a majority of patients in most populations bear a limited number of relatively common mutant alleles. Founder effect and drift would require the presence of a large number of distinct founding populations, since PKU occurs in a large number of different ethnic groups. This mechanism has been proposed to account for the frequency of a deletion mutation among Yemenite Jews (Avigad et al. 1990) and for the M1V (John et al. 1989; Lyonnet et al. 1992) and R408W (John et al. 1990) alleles in French-Canadians. The data collected in the present study suggest that this mechanism may also have played an important role in determining the current distribution of the major PKU alleles in European populations, since five of the most predominant PAH mutations are each strongly associated with a single RFLP haplotypes. Heterozygote advantage remains a persistently attractive mechanism in accounting for the frequency of PKU. However, PKU is much less prevalent than other genetic disorders, such as the globin disorders, where this mechanism has also been invoked. Thus, if heterozygote advantage does exist in PKU, its effect may be too small to detect or define. Of course, it is also possible that such selection has existed in the past but is no longer active. In either case, this selective advantage must have existed in regions where there were significantly different climatic, cultural, and dietary conditions, since PKU is present at a relatively high frequency not only in European populations but in some Asian and Middle Eastern populations as well (Liu and Zuo 1986; Daiger et al. 1989*b*).

The frequency distributions of these major mutant PAH alleles are similar to those observed in studies of other genetic markers in European populations (Menozzi et al. 1978; Sokal et al. 1991). For example, Menozzi et al. (1978) demonstrated several principal axes along which the frequencies of several blood group markers varied. One major axis ran from southeast to northwest, and those authors proposed that this cline reflected the gene flow that accompanied the introduction of early agricultural methods 5,000–10,000 years ago. This frequency gradient is roughly similar to that observed for the IVS10nt546 mutant allele in the present study, which had a high frequency in Mediterranean and Middle Eastern populations and which had diminished frequencies in northern and western populations. A second cline, running from east to west, was observed by both Menozzi et al. (1978) and by Sokal et al. (1991) in a study of the ABO protein markers among Europeans. These gradients are very similar to that observed in the present study for the R408W mutation on haplotype 2 chromosomes. The underlying causes of this east-west cline are less clear but may reflect the migrations of early Slavic or Germanic peoples in the middle of the first millenium A.D. The frequency distribution of the IVS12nt1 PAH mutation does not conform to any of the patterns previously observed by Menozzi et al. or Sokal et al. but does bear a strong resemblance to that observed for the ΔF508 CF mutation in studies compiled and reported by DeVoto and co-workers (European Working Group on CF Genetics 1990), suggesting that human migration rather than selection may have been the most important factor in the spread of these alleles.

One remaining question is the relative time frame when these PAH mutations first occurred in these putative founding populations. While PKU for decades was considered primarily a Caucasian disorder, recent newborn screening studies for PKU in China have established an incidence (about 1 in 16,500 births; Liu and Zuo 1986) similar to that in Caucasian populations. RFLP haplotype analysis of the PAH locus among Orientals has indicated a distribution of haplotypes distinct from that of Caucasians, suggesting a fundamental difference in the genetic basis of PKU in

the two races. This conclusion is further supported by a number of studies of PAH mutations in Chinese PKU patients (Tsai et al. 1990; Wang et al. 1989, 1991*a*, 1991*b*, 1991*c*, 1992), which indicate that only 2 of the 20 or so PAH mutations that account for over 70% of all mutant alleles in Orientals are present in both races. The two exceptions, the R261X (Shirahase et al. 1991) and R408W (Tsai et al. 1990) mutations, occur on different haplotype backgrounds and are likely to be the result of recurrent mutation. It can thus be concluded that at least a most, if not all, PAH mutations have occurred after the divergence of the caucasoid and the mongoloid peoples. Further refinement of the time frames of the mutational events that occurred in these founding groups will require much more detailed observations, not only from studies of PAH alleles but also from a large number of independent markers.

Acknowledgment

We wish to express our sincere gratitude to the many dedicated individuals in the European countries who contributed to the collection of blood samples from PKU patients and their kindreds, which made this study possible. We also thank Drs. Stephen Daiger and Luca Cavalli-Sforza for insightful discussions and Dr. Charles Scriver for his constant inspiration. The work was supported in part by NIH grant HD-17711 to S.L.C.W., who is also an Investigator of the Howard Hughes Medical Institute.

References

Abadie V, Lyonnet S, Maurin N, Berthelon M, Caillaud C, Giraud F, Mattei JF, et al (1989) CpG dinucleotides are mutation hot spots in phenylketonuria. Genomics 5:936–939

Aoki K, Wada Y (1988) Outcome of the patients detected by newborn screening in Japan. Acta Paediatr 30:429–434

Apold J, Eiken HG, Odland E, Fredriksen Å, Bakken A, Lorens JB, Boman H (1990) A termination mutation prevalent in Norwegian haplotype 7 phenylketonuria genes. Am J Hum Genet 47:1002–1007

Aulehla-Scholz C, Vorgerd M, Sautter E, Leupold D, Mahlmann R, Ullrich K, Olek K, et al (1988) Phenylketonuria: distribution of DNA diagnostic patterns in German families. Hum Genet 78:353–355

Avigad S, Cohen BE, Bauer R, Schwartz G, Frydman M, Woo SLC (1990) A single origin of phenylketonuria in Yemenite Jews. Nature 344:168–170

Bickel H, Bachmann C, Beckers R (1981) Neonatal mass screening for metabolic disorders. Eur J Pediatr 137:133–139

Chakraborty, R (1986) Estimation of linkage disequilibrium from conditional haplotype data: application to β-globin gene cluster in American blacks. Genet Epidemiol 3:323–333

Chakraborty R, Lidsky AS, Daiger SP, Guttler F, Sullivan S, DiLella AG, Woo SLC (1987) Polymorphic DNA haplotypes at the human phenylalanine hydroxylase locus and their relationship with phenylketonuria. Hum Genet 76: 40–46

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

Daiger SP, Chakraborty R, Reed L, Fekete G, Schuler D, Berenssi G, Nasz I, et al (1989*a*) Polymorphic DNA haplotypes at the phenylalanine hydroxylase (PAH) locus in European families with phenylketonuria (PKU). Am J Hum Genet 45:310–318

Daiger SP, Reed L, Huang S-S, Zeng Y-T, Wang T, Lo WHY, Okano Y, et al (1989*b*) Polymorphic DNA haplotypes at the phenylalanine hydroxylase (PAH) locus in Asian families with phenylketonuria (PKU). Am J Hum Genet 45:319–324

Dasovich M, Konecki D, Lichter-Konecki U, Eisensmith RC, Güttler F, Naughton E, Mullins C, et al (1991) Molecular characterization of a PKU allele prevalent in southern Europe and Ireland. Somat Cell Mol Genet 17:303–309

Dianzani I, Devoto M, Camaschella C, Saglio G, Gerrero GB, Cerone R, Romano C, et al (1990) Haplotype distribution and molecular defects at the phenylalanine hydroxylase locus in Italy. Hum Genet 86:69–72

DiLella AG, Huang WM, Woo SLC (1988) Screening for phenylketonuria mutations by DNA amplification with the polymerase chain reaction. Lancet 1:497–499

DiLella AG, Marvit J, Brayton K, Woo SLC (1987) An amino-acid substitution involved in phenylketonuria is in linkage disequilibrium with DNA haplotype 2. Nature 327:333–336

DiLella AG, Marvit J, Lidsky AS, Guttler F, Woo SLC (1986) Tight linkage between a splicing mutation and a specific DNA haplotype in phenylketonuria. Nature 322: 799–803

Dworniczak B, Aulehla-Scholz C, Horst J (1989) Phenylketonuria: detection of a frequent haplotype 4 allele mutation. Hum Genet 84:95–96

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

Dworniczak B, Kalaydjieva L, Aulehla-Scholz C, Ullrich K, Kremensky I, Radeva B, Horst J (1991*b*) Recurrent nonsense mutation in exon 7 of the phenylalanine hydroxylase gene. Hum Genet 87:731–733

Eisensmith RC, Woo SLC (1991) Phenylketonuria and the phenylalanine hydroxylase gene. Mol Biol Med 8:3–18

——— (1992*a*) Molecular basis of phenylketonuria and related hyperphenylalaninemias: mutations and polymorphisms in the human phenylalanine hydroxylase gene. Hum Mutat 1:13–23

——— (1992*b*) Updated listing of haplotypes at the human phenylalanine hydroxylase (PAH) locus. Am J Hum Genet (in press)

European Working Group on CF Genetics (1990) Gradient of distribution in Europe of the major CF mutation and of its associated haplotype. Hum Genet 85:436–441

Fölling A (1934) Über Ausscheidung von Phenylbrenztraubensäure in den Harn als Stoffwechselanomalie in Verbindung mit Imbezillität. Z Physiol Chem 227:169–176

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

Herrmann F, Wulff K, Wehnert M, Siedlitz G, Güttler F (1988) Haplotype analysis of classical and mild phenotype of phenylketonuria in the German Democratic Republic. Clin Genet 34:176–180

Hill WG, Robertson A (1968) Linkage disequilibrium in finite populations. Theor Appl Genet 38:226–231

Hofman KJ, Antonarakis SE, Missiou-Tsangaraki S, Boehm CD, Valle D (1989) Phenylketonuria in the Greek population. Mol Biol Med 6:245–250

John SWM, Rozen R, Laframboise R, Laberge C, Scriver CR (1989) Novel PKU mutation on haplotype 2 in French-Canadians. Am J Hum Genet 45:905–909

John SWM, Rozen R, Scriver 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

Kalaydjieva L, Dworniczak B, Aulehla-Scholz C, Devoto M, Romeo G, Stuhrmann M, Horst J (1991) Phenylketonuria mutation in southern Europeans. Lancet 337:865

Kalaydjieva L, Dworniczak B, Aulehla-Scholz C, Kremensky J, Bronzova J, Eigel A, Horst J (1990) Classical phenylketonuria in Bulgaria: RFLP haplotypes and frequency of the major mutations. J Med Genet 27:742–745

Konecki DS, Lichter-Konecki U (1991) The phenylketonuria locus: current knowledge about alleles and mutations of the phenylalanine hydroxylase gene in various populations. Hum Genet 87:377–388

Lewontin RC (1964) The interaction of selection and linkage. I. General considerations: heterotic models. Genetics 49:49–67

Lichter-Konecki U, Schlotter M, Konecki DS, Labeit S, Woo SLC, Trefz FK (1988) Linkage disequilibrium between mutation and RFLP haplotype at the phenylalanine hydroxylase locus in the German population. Hum Genet 78:347–352

Lichter-Konecki U, Schlotter M, Yaylak C, Özgüç M, Çoskun T, Özalp I, Wendel U, et al (1989) DNA haplotype analysis at the phenylalanine hydroxylase locus in the Turkish population. Hum Genet 81:373–376

Lidsky AS, Ledley FD, DiLella AG, Kwok SCM, Daiger SP, Robson KJH, Woo SLC (1985) Extensive restriction site polymorphism at the human phenylalanine hydroxylase locus and application in prenatal diagnosis of phenylketonuria. Am J Hum Genet 37:619–634

Liu SR, Zuo QH (1986) Newborn screening for phenylketonuria in eleven districts. Chi Med J 99:113–118

Lyonnet S, Melle D, de Braekeleer M, Laframboise R, Rey F, John SWM, Berthelon M, et al (1992) Time and space clusters of the French-Canadian M1V phenylketonuria mutation in France. Am J Hum Genet 51:191–196

Menozzi P, Piazza A, Cavalli-Sforza L (1978) Synthetic maps of human gene frequencies in Europeans. Science 201:786–792

Okano Y, Wang T, Eisensmith RC, Steinmann B, Gitzelmann R, Woo SLC (1990) Missense mutations associated with RFLP haplotypes 1 and 4 of the human phenylalanine hydroxylase gene. Am J Hum Genet 46:18–25

Orkin SH, Kazazian HH (1984) The mutation and polymorphism of the human β-globin gene and its surrounding DNA. Annu Rev Genet 18:131–171

Özalp I, Coskun T, Ceyhan M, Tokol S, Oran O, Erdem G, Tekinalp G, et al (1986) Incidence of phenylketonuria and hyperphenylalaninemia in a sample of the newborn population. J Inherit Metab Dis 9, Suppl 2:237–239

Perez B, Desviat LR, Die M, Ugarte M (1992) Mutation analysis of phenylketonuria in Spain: prevalence of two Mediterranean mutations. Hum Genet 89:341–342

Rey F, Berthelon M, Caillaud C, Lyonnet S, Abadie V, Blandin-Savoja F, Feingold J, et al (1988) Clinical and molecular heterogeneity of phenylalanine hydroxylase deficiencies in France. Am J Hum Genet 43:914–921

Riess O, Michael A, Speer A, Meiske W, Cobet G, Coutelle C (1988) Linkage disequilibrium between RFLP haplotype 2 and the affected PAH allele in PKU families from the Berlin area of the German Democratic Republic. Hum Genet 78:343–346

Romano V, Bosco P, Chiavetta V, Pitronaci L, Fasulo G, Mollica F, Giovannini M, et al. Geographical distribution of phenylalanine hydroxylase alleles in Sicily. Brain Dys (in press)

Scriver CR, John SWM, Rozen R, Eisensmith R, Woo SLC. Associations between populations, PKU mutations and RFLP haplotypes at the PAH locus: an overview. Brain Dys (in press)

Scriver CR, Kaufman S, Woo SLC (1989) The hyperphenylalaninemias. In: Scriver, Beaudet, Sly, Valle (eds) The metabolic basis of inherited disease, 6th ed. McGraw-Hill, New York, pp 495–546

Shirahase W, Oya N, Shimada M (1991) A new single base substitution in a Japanese phenylketonuria (PKU) patient. Brain Dev 13:283–284

Sokal RR, Oden NL, Wilson C (1991) Genetic evidence for the spread of agriculture in Europe by demic diffusion. Nature 351:143–145

Stuhrmann M, Riess O, Mönch E, Kurdoglu G (1989) Haplotype analysis of the phenylalanine hydroxylase gene in Turkish phenylketonuria families. Clin Genet 36:117–121

Sullivan SE, Moore SD, Connor JM, King M, Cockburn F, Steinmann B, Gitzelmann R, et al (1989) Haplotype distribution of the human phenylalanine hydroxylase locus in Scotland and Switzerland. Am J Hum Genet 44: 652–659

Svensson E, von Döbeln U, Hagenfeldt L (1991) Polymorphic DNA haplotypes at the phenylalanine hydroxylase locus and their relation to phenotype in Swedish phenylketonuria families. Hum Genet 87:11–17

Tsai T-F, Hsiao K-J, Su T-S (1990) Phenylketonuria mutation in Chinese haplotype 44 identical with haplotype 2 mutation in northern-European Caucasians. Hum Genet 84:409–411

Wang T, Okano Y, Eisensmith RC, Harvey ML, Lo WHY, Yuan LF, Huang SZ, et al (1991*a*) Founder effect of a prevalent PKU mutation in the Oriental population. Proc Natl Acad Sci USA 88:2146–2150

Wang T, Okano Y, Eisensmith R, Huang S-Z, Zeng Y-T, Lo WHY, Woo SLC (1989) Molecular genetics of phenylketonuria in Orientals: linkage disequilibrium between a termination mutation and haplotype 4 of the phenylalanine hydroxylase gene. Am J Hum Genet 45:675–680

Wang T, Okano Y, Eisensmith RC, Lo WHY, Huang S-Z, Zeng Y-T, Liu S-R, et al (1991*b*) Missense mutations prevalent in Orientals with phenylketonuria: molecular characterization and clinical implications. Genomics 10: 449–456

Wang T, Okano Y, Eisensmith RC, Lo WHY, Huang S-Z, Zeng Y-T, Woo SLC (1991*c*) Identification of a novel phenylketonuria (PKU) mutation in the Chinese: further evidence for multiple origins of PKU in Asia. Am J Hum Genet 48:628–630

Wang Y, Okano Y, Eisensmith RC, Lo WHY, Huang S-Z, Zeng Y-T, Yuan L-F, et al (1992) Identification of three novel PKU mutations among Chinese: evidence for recombination or recurrent mutation at the PAH locus. Genomics 13:230–231

Woo SLC (1988) Collation of RFLP haplotypes at the human phenylalanine hydroxylase (PAH) locus. Am J Hum Genet 43:781–783

Woo SLC, Lidsky A, Chandra T, Güttler F, Robson K (1983) Cloned human phenylalanine hydroxylase gene allows prenatal diagnosis and carrier detection of classical phenylketonuria. Nature 306:151–155

Zygulska M, Eigel A, Aulehla-Scholz C, Pietrzyk JJ, Horst J (1991) Molecular analysis of PKU haplotypes in the population of southern Poland. Hum Genet 86:292–294