Am. J. Hum. Genet. 44:382–387, 1989

Phenylalanine Hydroxylase Gene Haplotypes in Polynesians: Evolutionary Origins and Absence of Alleles Associated with Severe Phenylketonuria

M. Hertzberg,* K. Jahromi,* V. Ferguson,* H. H. M. Dahl,† J. Mercer,† K. N. P. Mickleson,‡ and R. J. Trent*

*Molecular Biology Laboratory, Clinical Immunology Research Centre, University of Sydney, Sydney; †Murdoch Research Institute, Royal Children's Hospital, Parkville, Victoria, Australia; and ‡Department of Paediatrics, Middlemore Hospital, Auckland

Summary

A total of 630 haplotypes for the phenylalanine hydroxylase (PAH) gene locus were established in five groups of Polynesians comprising Samoans, Tongans, Cook Islanders, Maori, and Niueans. Considerable genetic continuity was demonstrated between these widely dispersed populations, since three common haplotypes (4, 1, and 7) constituted over 95% of alleles. A control group of individuals from Southeast Asia shared the same major haplotypes, 4, 1, and 7, with Polynesians. These data provide further support for the theories of genetic homogeneity and of Asian affinities of the Polynesian precursor populations. The absence of severe phenylketonuria (PKU) in both Polynesians and Southeast Asians is consistent with the lack of PAH haplotypes 2 and 3, on which the severe PKU mutants have arisen among Caucasians.

Introduction

The phenylalanine hydroxylase (PAH) gene is located on chromosome 12q22-q24.1 (Lidsky et al. 1985*a*). Occupying a region 90 kb in length, it consists of 13 exons and codes for a protein of 451 amino acids (Kwok et al. 1985). Severe deficiency of this enzyme results in classical phenylketonuria (PKU), an autosomal recessive disorder with a prevalence of one in 11,000 Caucasians (Scriver and Clow 1980). Recently, full-length PAH cDNA clones have been isolated (Kwok et al. 1985; Speer et al. 1986). Eight RFLPs at the PAH locus have been detected, which has enabled the construction of DNA haplotypes (Lidsky et al. 1985*a*). The eight polymorphic restriction sites exist in significant linkage disequilibrium, since only 12 of a possible 1,536 haplotypes have been found in the Danish population (Chakraborty et al. 1987). Strong linkage disequilibrium between PKU mutants and RFLP haplotypes has enabled the use of these haplotypes in prenatal and carrier detection of PKU (Lidsky et al. 1985*b*; Daiger et al. 1986).

In the present study, RFLP haplotypes for the PAH locus were established in five Polynesian groups as well as in a population of Southeast Asians. Significant differences are demonstrated by both the above when compared with northern Europeans, to date the only other extensively studied population (Chakraborty et al. 1987; Lichter-Konecki et al. 1988). We confirm that limited diversity of haplotypes is exhibited by the widely dispersed Polynesian island groups, which is consistent with the theories of a relatively homogeneous founding population in Polynesia. Furthermore, Polynesians share major PAH haplotypes with Southeast Asians but lack those haplotypes associated with classical PKU.

Subjects and Methods

Subjects

A total of 347 Polynesian subjects representing Samoans, Tongans, Cook Islanders, Maori, and Niueans were studied. These included 30 Samoan and two Maori families, each consisting of the two parents and at least one child. A further 251 random individuals were sub-

Received September 13, 1988; revision received November 7, 1988.

Address for correspondence and reprints: Dr. M. Hertzberg, Molecular Biology Laboratory, Clinical Immunology Research Centre, University of Sydney, Sydney, NSW 2006, Australia.

0002-9297/89/4403-0011$02.00

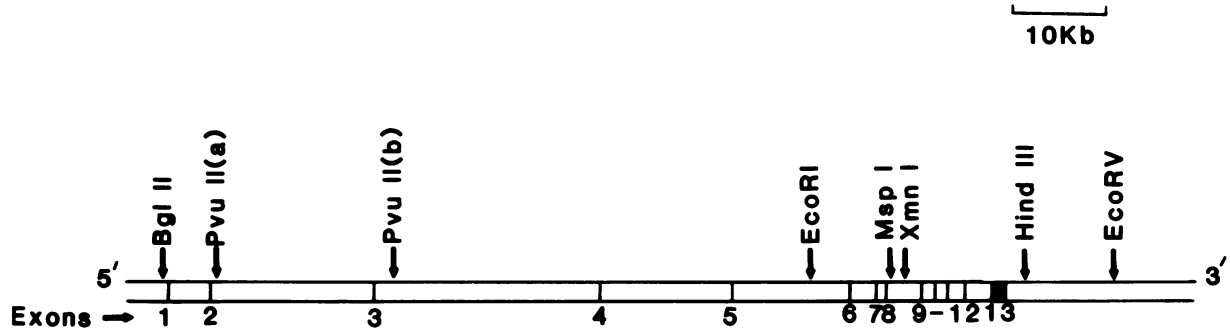

Figure 1 PAH gene with polymorphic restriction sites as described in the text

sequently analyzed. Samples were obtained from: (1) families living in rural villages on the island of Savaii in Western Samoa; (2) random cord bloods derived from Polynesian newborns at the Middlemore Hospital, Auckland; and (3) adult Polynesians living in Auckland. Ethnic origins were established by direct questioning of adults sampled or from mothers of newborns. A second sample group, comprising Southeast Asians, was also studied. It included eight families and 21 random individuals, all of whom were living in Sydney, but who had originated from Hong Kong, Indochina, or the Philippines. Each family consisted of two parents and one child.

DNA Mapping

Blood samples were collected in heparin and frozen until DNA was prepared by standard phenol/chloroform extraction (Old and Higgs 1983). Eight-microgram aliquots of DNA were digested with the seven restriction enzymes *Bgl*II, *Pvu*II, *Eco*RI, *Msp*I, *Xmn*I, *Hind*III, and *Eco*RV (fig. 1). DNA was separated on horizontal agarose gels (0.7%–1.0%) and transferred onto nylon membranes (Hybond,™ Amersham) with 20 × SSC. The hybridization probe was a 2.3-kb human cDNA insert encoding the entire PAH mRNA, including 40 bp of 5′- and 850 bp of 3′-untranslated regions, and was inserted into the *Eco*RI sites of pUC9 (Speer et al. 1986). The cDNA PAH probe was labeled with ^{32}P dCTP by using either the nick-translation kit or the multiprime DNA labeling kit (Amersham). Hybridization was performed in a solution containing 6 × SCC, 5 × Denhardt's, 0.5% SDS, 20 μg salmon sperm DNA/ml, and 5% dextran sulfate at 65°C for 18 h (Old and Higgs 1983). Washing of membranes was carried out at 65°C in 0.1 × SSC for 1.5 h, followed by autoradiography using Cronex Quanta III™ intensifying screens.

Results

Derivation of PAH Haplotypes

Six hundred thirty restriction-enzyme haplotypes for the PAH locus were determined for individuals who originated from five Polynesian islands. Initially, haplotypes were established from family studies (table 1). Of the 32 families included in the study, 26 were unequivocally informative, providing a total of 104 alleles. In establishing haplotypes in one-child families, the absence of recombination was assumed. Haplotype assignment in random individuals homozygous for polymorphisms accounted for an additional 214 alleles. Subsequently, it was possible to classify PAH alleles for the remaining six families and random heterozygous individuals in the vast majority of cases. A further 299 haplotypes were obtained in this manner, leaving 13 alleles that could not be confidently assigned. Haplotype distributions were comparable among the three groups (table 1).

PAH Haplotypes in Polynesians

Haplotypes 4, 1, and 7 accounted for the vast majority of PAH alleles in each of the five Polynesian populations (table 2). Overall, haplotype 4 represented 58% of alleles, ranging from 53% in Cook Islanders to 69% in Niueans. Haplotypes 1 and 7 were present at frequencies of 21% (between extremes of 20% and 26%) and 16% (between extremes of 5% and 23%), respectively. Haplotype 3 was found infrequently in Polynesians, as was haplotype 6. Additionally, two new haplotypes, 20 (− − + + + − − −) and 31 (− + − + − + + −) (Woo 1988), were detected from the informative Samoan families (table 1). Haplotype 20 is characterized by the presence of the *Pvu*IIb restriction site and differs from haplotype 6 at the *Bgl*II polymorphic site. Haplotype 31 differs from haplotype 4 at the *Eco*RV site. Thirteen

Table 1

RFLP Haplotypes at the PAH Locus

Type	B	P(a)	P(b)	R	M	X	H	V	No. (%) of: Polynesians Families	Polynesians Homozygotes	Polynesians Heterozygotes	Southeast Asians Families	Southeast Asians Homozygotes	Southeast Asians Heterozygotes
1	−	+	−	−	+	−	−	−	22 (21)	24 (11)	89 (29)	2 (6)	2 (11)	2 (8)
2	−	+	−	−	+	−	+	+	0	0	0	0	0	0
3	−	+	−	+	−	+	−	−	1 (1)	0	4 (1)	0	0	1 (4)
4	−	+	−	+	−	+	+	+	63 (61)	174 (81)	129 (41)	27 (84)	14 (78)	12 (50)
5	+	−	+	+	+	−	−	+	0	0	0	0	0	0
6	+	−	+	+	+	−	−	−	0	0	6 (2)	0	0	0
7	+	−	−	+	−	+	−	−	16 (15)	16 (8)	68 (22)	1 (3)	2 (11)	5 (21)
20	−	−	+	+	+	−	−	−	1 (1)	0	2 (1)	0	0	0
31	−	+	−	+	−	+	+	−	1 (1)	0	1 (<1)	0	0	0
44	+	−	−	+	−	+	+	+	0	0	0	2 (6)	0	3 (13)
Unclassified									0	0	13 (4)	0	0	1 (4)
Total nos. tested									104	214	312	32	18	24

NOTE.—Polymorphic restriction sites are as described in the text and in fig. 1: B = *Bgl*II, P(a) = *Pvu*II (a), P(b) = *Pvu*II (b), R = *Eco*RI, M = *Msp*I, X = *Xmn*I, H = *Hin*dIII, V = *Eco*RV. Haploytpes are classified according to Woo (1988).

alleles from heterozygous individuals remained unclassified.

PAH Haplotypes in Southeast Asians

A strategy similar to that described above was applied to individuals of Southeast Asian origin (tables 1, 2). All eight families were unequivocally informative, providing a total of 32 haplotypes. An additional 18 haplotypes were obtained in individuals homozygous for polymorphisms. Subsequently, PAH alleles (23) could be assigned for heterozygous individuals in nearly all cases. Approximately 72% of the haplotypes present in Southeast Asians were type 4. PAH alleles 1 and 7 were detected at frequencies of 8% and 11%, respectively. A new haplotype 44 (+−−+−+++), differing from haplotype 4 at the *Bgl*II and *Pvu*IIa polymorphic

Table 2

RFLP Haplotypes at the PAH Locus in Polynesians and Southeast Asians[a]

Haplotype	No. (%) of: Samoans	Tongans	Cook Islanders	Maori	Niueans	Total Polynesians	Asians
1	37 (20)	18 (24)	43 (20)	20 (26)	17 (23)	135 (21)	6 (8)
2	0	0	0	0	0	0	0
3	3 (2)	0	0	2 (3)	0	5 (1)	1 (1)
4	116 (62)	44 (58)	114 (53)	41 (53)	51 (69)	366 (58)	53 (72)
5	0	0	0	0	0	0	0
6	1 (1)	1 (1)	4 (2)	0	0	5 (1)	0
7	23 (12)	10 (13)	50 (23)	13 (17)	4 (5)	100 (16)	8 (11)
20	3 (2)	0	0	0	0	3 (<1)	0
31	2 (1)	0	0	0	0	2 (<1)	0
44	0	0	0	0	0	0	5 (7)
Unclassified	3 (2)	3 (4)	3 (1)	2 (3)	2 (3)	13 (2)	1 (1)
Total	188	76	214	78	74	630	74

[a] Results are total numbers obtained for families, homozygotes, and heterozygotes.

sites, was determined from family studies. This novel haplotype was found at a frequency of 7%.

Discussion

RFLP haplotypes, such as those for the α- and β-globin gene clusters, represent valuable markers with which to examine the genetic diversity and evolutionary relationships of modern populations (Higgs et al. 1986; Wainscoat et al. 1986). The PAH gene also provides a very polymorphic marker for population studies (Lidsky et al. 1985*a*). The degree of heterozygosity at this locus is 88%, which is comparable to that described for the globin gene clusters. Despite this, polymorphisms present at the PAH locus exist in considerable linkage disequilibrium, since populations studied to date exhibit only a small number of haplotype combinations. For example, 12 common RFLP haplotypes defined at the PAH locus were found in a Danish population (Chakraborty et al. 1987).

In the present study, PAH haplotypes were derived initially from families. Additional haplotypes were obtained from random samples that were either homozygous for the polymorphisms or heterozygotes for which haplotypes could be assigned. Allele distributions were comparable among the three groups. Thirteen haplotypes from Polynesians remained unclassified. Hence, some rare or novel normal PAH alleles may have been undetected. However, since these represented less than 2% of the total alleles, any ascertainment bias would be extremely small. Haplotype 4 represents a significant majority of PAH alleles found in Polynesians. Overall, it accounted for 58% of PAH haplotypes and, in particular, represented 69% of alleles found in Niueans. While haplotype 4 is also found in northern Europeans, it is present at significantly lower frequencies of 19%–32% (table 3). The relative paucity of haplotype 3 (<1%) and apparent absence of haplotypes 2 and 5 in Polynesians also make them significantly different from northern Europeans, in whom these three haplotypes together comprise between 19% and 25% of normal alleles (Chakraborty et al. 1987; Lichter-Konecki et al. 1988). PAH haplotype 1 was detected in Polynesians at a frequency of 21%. This contrasts with the Danes, in whom haplotype 1 represents 35% of normal PAH alleles (Chakraborty et al. 1987). Additionally, two recently described haplotypes, 20 (−−+++−−−) and 31 (−+−+−+++−), were found in Samoan families. It is unlikely that the haplotype distributions present in Polynesians represent European admixture. Each of the five independent island groups exhibited very similar frequencies of alleles. Furthermore, globin and HLA genetic analyses (Trent et al. 1986; Serjeantson et al. 1987), together with mitochondrial DNA data (our own unpublished observations), confirm a paucity of European admixture in our sample population. In fact, one of the striking features to emerge from this study is the relative homogeneity of Polynesian populations. Each of the island groups exhibited limited diversity of their PAH alleles, in which three major haplotypes (4, 1, and 7) predominated. All population groups shared the above haplotypes, differing in their relative proportions of each. These findings are consistent with α-globin gene haplotype data, which similarly show limited diversity of alleles in Polynesians (Hertzberg et al. 1988), since six α-globin haplotypes represent over 96% of normal alleles. Taken together, these findings support the theories of a common, relatively small founding population in Polynesia.

The presence of haplotypes 4, 1, and 7 in all populations examined to date (including Caucasians, Southeast Asians, and Polynesians) suggests an origin for these alleles that predates divergence of the races. Haplotypes that have arisen following this divergence, including novel alleles 20, 31, and 44 in the present study, may have resulted either from chromosomal crossover within the common haplotypes or by mutation at single sites. For example, haplotype 44 in Southeast Asians could have arisen from DNA crossover involving haplotypes 4 and 7 between *Pvu*II and *Eco*RI restriction sites. In contrast, haplotype 31 in Polynesians may have resulted from a single nucleotide mutation at the *Eco*RV site on haplotype 4. Absence of haplotypes 2 and 5 in both Polynesians and Southeast Asians may reflect a loss of

Table 3

Normal PAH Haploytpes in Polynesians, Southeast Asians, and Nothern Europeans

Haplotype	Polynesians (%)	Asians (%)	Northern Europeans (%)	
			I[a]	II[b]
1	21	8	35	25
2	0	0	3	10
3	<1	1	5	5
4	58	72	32	19
5	0	0	11	10
6	1	0	0	2
7	16	11	11	16
Other	3	9	3	13

[a] Derived from Chakraborty et al. (1987).
[b] Derived from Lichter-Konecki et al. (1988).

these alleles after separation of the races. Alternatively, types 2 and 5 may represent new alleles that have arisen subsequently in Europeans. It is unlikely that these haplotype distributions result from selection, since, in the absence of classical PKU, site polymorphisms would appear to be neutral.

Genetic data, such as HLA markers, have provided conflicting evidence of the evolutionary relationships of Polynesians and Southeast Asians. Several Mongoloid markers, including HLA-A2B40 and HLA-A9Bw22 and some class II DNA patterns, are shared by both groups (Serjeantson et al. 1982; Kohonen-Corish and Serjeantson 1986). Conversely, there are a number of HLA genes present in Orientals but absent in Polynesians and vice versa (Serjeantson et al. 1987). The triplicated ζ-globin gene rearrangement (ζζζ/) is present in both Polynesians and Southeast Asians, at frequencies of 22% and 9%, respectively (Higgs et al. 1986; Trent et al. 1986). In each population, this rearrangement is associated exclusively with the same α-globin gene haplotype, suggesting that it has arisen once only during evolution (Hertzberg et al. 1988). In the present study, genetic similarities between Polynesians and Southeast Asians are further seen in RFLP haplotypes at the PAH locus. Together, haplotypes 4, 1, and 7 accounted for the vast majority of alleles present in both groups of peoples (95% and 92%, respectively). In particular, haplotype 4 constituted 58% and 72% of PAH alleles in Polynesians and Southeast Asians, respectively, while being found in only 19%–32% of Europeans. Similarly, both populations exhibited a relative paucity of haplotype 3, together with an absence of haplotypes 2 and 5. Therefore, PAH RFLP data lend support to the theories of a genetic affinity between Southeast Asians and the ultimate Polynesian and Lapita forebears (reviewed in Bellwood 1987).

Phenylketonuria is most prevalent among Celtic-derived people and central Europeans, while it is extremely uncommon in Southeast Asians (Scriver and Clow 1980). To date, it has been detected in one Polynesian newborn, and in this case there was a European parent (K.N.P.M., unpublished data). Recently, two PKU mutations have been characterized in Caucasians, and both of these result in the severe form of the disease (DiLella et al. 1986, 1987; Guttler et al. 1987). Each mutation is tightly linked with haplotypes 2 and 3. Together, these mutations represent 58% of total PKU chromosomes in the northern European population (DiLella et al. 1988). DiLella et al. (1986) have suggested that each of the above PKU mutations may have arisen only once throughout evolution — and that it may have spread by subsequent founder effect in Europeans. Haplotypes 2 and 3 represent approximately 8% of the normal alleles present in the latter population (DiLella et al. 1987). In contrast, Polynesians and Southeast Asians have a significantly lower frequency of haplotype 3 (<1%) as well as an absence of haplotype 2. Therefore, it is significant that both Polynesians and Southeast Asians, in whom PAH haplotypes 2 and 3 are virtually absent, also appear to lack the severe PKU mutations characterized previously in Europeans. The presence of these PKU mutants in European populations suggests an origin subsequent to separation of the races. Determination of whether the rare cases of mild PKU in Southeast Asians are due to the identical mutations carried on the same haplotypes as those present in northern Europeans awaits further analysis.

Acknowledgments

We wish to thank Dr. W. Vermeulen, Department of Health, Apia, Western Samoa, for his cooperation in enabling us to collect samples. We are grateful to Dr. S. L. C. Woo for providing us with the collated PAH haplotype data. This study was supported by the National Health and Medical Research Council of Australia; M. H. was supported by a Faculty of Medicine scholarship, University of Sydney, N.S.W.

References

Bellwood, P., ed. 1987. The polynesians: prehistory of an island people. Thames & Hudson, London.

Chakraborty, R., A. S. Lidsky, S. P. Daiger, F. Güttler, S. Sullivan, A. G. DiLella, and S. L. C. Woo. 1987. Polymorphic DNA haplotypes at the human phenylalanine hydroxylase locus and their relationship with phenylketonuria. Hum. Genet. **76**:40–46.

Daiger, S. P., R. Chakraborty, F. Güttler, A. S. Lidsky, R. Koch, and S. L. C. Woo. 1986. Polymorphic DNA haplotypes at the phenylalanine hydroxylase locus in prenatal diagnosis of phenylketonuria. Lancet **1**:229–232.

DiLella, A. G., J. Marvit, K. Brayton, and S. L. C. Woo. 1987. An amino-acid substitution involved in phenylketonuria is in linkage disequilibrium with DNA haplotype 2. Nature **327**:333–336.

DiLella, A. G., J. Marvit, A. S. Lidsky, F. Güttler, and S. L. C. Woo. 1986. Tight linkage between a splicing mutation and a specific DNA haplotype in phenylketonuria. Nature **322**:799–803.

DiLella, A. G., Whei-Mei Huang, and S. L. C. Woo. 1988. Screening for phenylketonuria mutations by DNA amplification with the polymerase chain reaction. Lancet **1**:497–499.

Güttler, F., F. D. Ledley, A. S. Lidsky, A. G. DiLella, S. E. Sullivan, and S. L. C. Woo. 1987. Correlation between polymorphic DNA haplotypes at phenylalanine hydroxylase locus and clinical phenotypes of phenylketonuria. J. Pediatr. **110**:68–71.

Hertzberg, M. S., K. N. P. Mickleson, and R. J. Trent. 1988. α-Globin gene haplotypes in Polynesians: their relationship to population groups and gene rearrangements. Am. J. Hum. Genet. **43**:971–977.

Higgs, D. R., J. S. Wainscoat, J. Flint, A. V. S. Hill, S. L. Thein, R. D. Nicholls, H. Teal, H. Ayyub, T. E. A. Peto, A. G. Falusi, A. P. Jarman, J. B. Clegg, and D. J. Weatherall. 1986. Analysis of the human α-globin gene cluster reveals a highly informative genetic locus. Proc. Natl. Acad. Sci. USA. **83**:5165–5169.

Kohonen-Corish, M. R. J., and S. W. Serjeantson. 1986. RFLP analysis of HLA-DR and -DQ genes and their linkage relationships in the Pacific. Am. J. Hum. Genet. **39**:751–762.

Kwok, S. C., F. D. Ledley, A. G. DiLella, K. J. H. Robson, and S. L. C. Woo. 1985. Nucleotide sequence of a full-length complementary DNA clone and amino acid sequence of human phenylalanine hydroxylase. Biochemistry **24**:556–561.

Lichter-Konecki, U., M. Schlotter, D. S. Konecki, S. Labeit, S. L. C. Woo, and F. K. Trefz. 1988. Linkage disequilibrium between mutation and RFLP haplotype at the phenylalanine hydroxylase locus in the German population. Hum. Genet. **78**:347–352.

Lidsky, A. S., F. D. Ledley, A. G. DiLella, S. C. M. Kwok, S. P. Daiger, K. J. H. Robson, and S. L. C. Woo. 1985*a*. Extensive restriction site polymorphism at the human phenylalanine hydroxylase locus and application in prenatal diagnosis of phenylketonuria. Am. J. Hum. Genet. **37**: 619–634.

Lidsky, A. S., F. Güttler, and S. L. C. Woo. 1985*b*. Prenatal diagnosis of classic phenylketonuria by DNA analysis. Lancet **1**:549–551.

Old, J. M., and D. R. Higgs. 1983. Gene analysis. Pp. 74–102 *in* D. J. Weatherall, ed. The thalassaemias: methods in haematology. Churchill Livingstone, Edinburgh.

Scriver, C. R., and C. L. Clow. 1980. Phenylketonuria: epitome of human biochemical genetics. N. Engl. J. Med. **303**:1394–1400.

Serjeantson, S. W., D. P. Ryan, and A. R. Thompson. 1982. The colonization of the Pacific: the story according to human leukocyte antigens. Am. J. Hum. Genet. **34**:904–918.

Serjeantson, S. W., B. S. White, E. C. Jazwinska, P. T. Yenchitsomanus, K. N. P. Mickleson, and R. J. Trent. 1987. HLA-DR and -DQ DNA genotyping in Polynesians and in Papua New Guinea highlanders. Hum. Immunol. **20**:145–153.

Speer, A., H. H. Dahl, O. Riess, G. Cobet, R. Hanke, R. G. H. Cotton, and Ch. Covtelle. 1986. Typing of families with classical phenylketonuria using three alleles of the *Hin*dIII linked restriction fragment polymorphism, detectable with a phenylalanine hydroxylase cDNA probe. Clin. Genet. **29**:491–495.

Trent, R. J., K. N. P. Mickleson, T. Wilkinson, J. Yakas, M. W. Dixon, P. J. Hill, and H. Kronenberg. 1986. Globin genes in Polynesians have many rearrangements including a recently described γγγγ/. Am. J. Hum. Genet. **39**:350–360.

Wainscoat, J. S., A. V. S. Hill, A. L. Boyce, J. Flint, M. Hernandez, S. L. Thein, J. M. Old, J. R. Lynch, A. G. Falusi, D. J. Weatherall, and J. B. Clegg. 1986. Evolutionary relationships of human populations from an analysis of nuclear DNA polymorphisms. Nature **319**:491–493.

Woo, S. L. C. 1988. Collation of RFLP haplotypes at the phenylalanine hydroxylase (PAH) locus. Am. J. Hum. Genet. **43**:781–783.