Molecular heterogeneity at the phenylalanine hydroxylase locus in the population of the south-west of England

L A Tyfield, M J Osborn, J B Holton

Abstract
The phenylalanine hydroxylase gene locus has been studied in 35 independent phenylketonuric families in the south-west of England using RFLP haplotype patterns and allele specific oligonucleotide probes. Haplotype 3 was the most common pattern on mutant chromosomes and there was strict linkage disequilibrium between this haplotype and the splice mutation in exon 12. The R408W mutation in exon 12 occurred on both haplotypes 1 and 2. The R126Q mutation in exon 7 was found only on a rare haplotype 28 pattern. No gene carried the R158Q mutation. More than 60% of mutant genes did not carry these four mutations which were originally described in other European populations. We suggest that the splice mutation arose as a single event and spread throughout northern Europe by population migration and admixture. In addition, we believe the haplotype/mutation associations seen in our population are a reflection of the mixed ancestry of the inhabitants of the British Isles.

The study of the phenylalanine hydroxylase (PAH) gene using haplotype patterns derived from restriction fragment length polymorphisms (RFLPs) originally uncovered extensive genetic variation at this locus both within and between populations.1-9 More recently, as an increasing number of mutations are defined,10-16 the combined use of RFLP haplotypes and allele specific oligonucleotide (ASO) probes is providing some insight into evolutionary aspects of the mutations that give rise to phenylketonuria (PKU) in different populations. Some mutations are found almost exclusively in individual populations.16 Others are found more widely and many of these have been reported in association with more than one haplotype.17-21

We have examined the PAH locus in the PKU population in the south-west of England using RFLP haplotype patterns and ASO probes. All PKU genes were examined for the presence of four mutations, which were originally found in other European populations10 11 15 in association with haplotypes 1, 2, 3, and 4. Our results are similar to those reported in other studies10 11 20 and lend support to the hypotheses that the mutation/haplotype associations that are seen on PKU chromosomes have arisen through hypermutability at specific sites, intragenic crossover, and gene conversion.17 20

Methods
The South Western Regional Health Authority covers five geographical counties in England: Gloucestershire, Avon, Somerset, Devon, and Cornwall. We analysed the PAH gene in 35 separate families in the south-west of England in whom there is a proband with PKU. Altogether 72 independent mutant genes were studied, one family having three affected subjects spanning three generations.22 Subjects with non-PKU hyperphenylalaninaemia were not included.

The full length cDNA probe PAH247, kindly supplied by Dr Savio Woo, Houston, Texas, was used to determine RFLP haplotypes.23 Haplotype numbers were the same as those used by Woo.24 Polymerase chain reaction (PCR) amplification, dot blots, and allele specific oligonucleotide probes were used to test mutant genes for the R261Q mutation in exon 7, the R408W mutation in exon 12, the splice donor mutation of exon 12, and the R158Q mutation in exon 5. These mutations were first reported on haplotypes 1,15 2,11 3,10 and 4,15 17 respectively, in other European populations. Details of the techniques have been described elsewhere.22

Results
HAPLOTYPE PATTERNS AT THE PAH LOCUS
Considerable genetic heterogeneity occurs at the PAH locus in both mutant and non-PKU chromosomes.
Molecular heterogeneity at the phenylalanine hydroxylase locus in the population of the south-west of England

Table 1  Characterisation of the phenylalanine hydroxylase gene using RFLP haplotype patterns.

<table>
<thead>
<tr>
<th></th>
<th>Mutant</th>
<th>Normal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total no of chromosomes</td>
<td>72</td>
<td>64</td>
</tr>
<tr>
<td>No of alleles assigned haplotype numbers</td>
<td>52 (72%)</td>
<td>45 (70%)</td>
</tr>
<tr>
<td>No of different haplotype patterns</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Most common haplotypes</td>
<td>1,2,3,4 (67%)</td>
<td>1,4,5 (64%)</td>
</tr>
</tbody>
</table>

Table 2  Showing the distribution of the more common haplotype patterns on mutant and non-mutant genes.

<table>
<thead>
<tr>
<th>Haplotype pattern</th>
<th>PKU alleles</th>
<th>Non-PKU alleles</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No %</td>
<td>No %</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>13</td>
<td>25</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>13</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>35</td>
<td>30</td>
</tr>
</tbody>
</table>

With about equal frequency on mutant and non-mutant chromosomes.

ASSOCIATION BETWEEN HAPLOTYPE PATTERNS AND SPECIFIC MUTATIONS

All 72 PKU genes were tested for each of four mutations associated with haplotypes 1, 2, 3, and 4 in other European populations. Although fewer than 40% of mutant genes carried these mutations, the splice mutation in exon 12 was the most common and there was linkage disequilibrium between a haplotype 3 pattern and this mutation (table 3). The splice mutation was also found in two unrelated children on a haplotype which has not been previously listed. This pattern is identical to a haplotype 3 except for the loss of the EcoRI recognition site. Four other mutant genes in three unrelated children also carried the splice mutation. The DNA of one of these children had degraded and so RFLP studies could not be completed; however, PCR amplification of the degraded DNA showed him to be homozygous for the splice mutation. The other two splice mutations almost certainly occurred on haplotype 3 patterns, but extensive RFLP heterozygosity in the two families precluded patterns being assigned with certainty.

The R408W mutation in exon 12 was found on haplotype 2 mutant alleles; however, more mutant haplotype 2 chromosomes did not carry this mutation. In addition, the mutation was found on some mutant chromosomes carrying a haplotype 1 pattern. One child who was homozygous for haplotype 1 was heterozygous for the R408W mutation and an unknown mutation. The point mutation was also found on chromosomes whose haplotype patterns could not be assigned. In all cases, because of the absence of HindIII and EcoRI recognition sites, it was certain that the mutations were not associated with a haplotype 2 pattern.

No mutant haplotype 1 allele carried the exon 7

Table 3  The association of four specific mutations with various haplotype patterns at the PAH locus. All mutant genes were tested for each mutation.

<table>
<thead>
<tr>
<th>Haplotype patterns</th>
<th>R261Q exon 7</th>
<th>R408W exon 12</th>
<th>Splice donor junction exon 12</th>
<th>R158Q exon 5</th>
<th>Other (not identified)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>13</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>Other (haplotype defined)</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>14</td>
</tr>
<tr>
<td>Other (haplotype not assigned*)</td>
<td>0</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>Total</td>
<td>1</td>
<td>8 (11%)</td>
<td>19 (26%)</td>
<td>0</td>
<td>44 (61%)</td>
</tr>
</tbody>
</table>

*Because of extensive heterogeneity within the family or DNA degradation.
R261Q mutation; however, this mutation was found in one child on a rare haplotype 28 pattern. No PKU gene carried the exon 5 R158Q mutation which has been associated with haplotype 4.

Discussion

The study of the PAH gene using RFLP haplotype patterns originally uncovered extensive genetic heterogeneity at this locus. Although most PKU chromosomes are associated with only four haplotypes, there are differences in the frequency distribution of haplotype patterns in different populations.\(^1\)\(^-\)\(^6\) The PKU families in the south-west of England are all Caucasian and are of European ancestry. In some respects the distribution of haplotype patterns is similar to the Danish\(^1\) and Scottish\(^6\) populations in that a haplotype 3 was found most frequently on mutant chromosomes. Unlike the Danish and other north European studies,\(^2\)\(^3\)\(^18\) however, haplotype 2 is relatively rare in our population.

Several specific mutations which give rise to hyperphenylalaninaemia have been defined.\(^10\)\(^-\)\(^16\) Although some are population specific, others are more widespread and several have been found to be associated with more than one haplotype pattern. Various authors have suggested that the mutation/haplotype associations that occur in different populations are evidence of hypermutability at specific sites, gene conversion, and intragenic crossover at the PAH locus.\(^17\)\(^20\)

The E280K mutation in exon 7, for example, which was originally described on a haplotype 38 in the Algerian population\(^15\) was also found on haplotypes 4 and 1 in a French\(^24\) and Danish family,\(^17\) respectively. This mutation was not found on several haplotype 1 chromosomes in our own population (unpublished observations). Haplotypes 38, 4, and 1 differ at several or all polymorphic sites, and since the linkage disequilibrium is exclusively maintained in each of the populations, it has been proposed that the same mutation arose at the same site as independent events on different haplotype backgrounds.\(^17\) Furthermore, a CpG dinucleotide is involved. CpG is a common site of methylation in mammalian DNA and spontaneous deamination of the dinucleotide has been implicated as a 'hot spot' for mutations, showing a rate of mutation 42-fold higher than that of random mutation.\(^25\)

The R408W mutation in exon 12 has also been found on different haplotype backgrounds in independent populations. In northern Europe, it was originally reported in linkage disequilibrium with haplotype 2; however, in the French Canadian population, it is associated with a haplotype 1 pattern. Intragenic crossover 5' to the HindIII site could change haplotype 2 to a haplotype 1; however, Okano et al\(^17\) have suggested that a mutation should appear on more than one haplotype in any single population if intragenic crossover is to be considered as the underlying cause. John et al\(^20\) proposed that independent mutational events are probably responsible for the appearance of the R408W mutation on each haplotype, particularly since a CpG dinucleotide is involved. The fact that haplotype 1 is common at the PAH locus and more than one mutation is associated with each of haplotypes 1 and 2 provides additional support to this hypothesis.

In our study the R408W mutation was found on both haplotypes 1 and 2. Although intragenic crossover could have occurred independently in our own population the observed associations are more likely to be a reflection of the mixed ancestry of the inhabitants of the British Isles and could represent independent legacies of invasions from the north and south 1000 to 2000 years ago.

It is interesting that one child is homozygous for the haplotype 1 pattern but is a compound heterozygote for the specific mutations which gave rise to PKU. This sort of heterogeneity could account for the diverse findings in early reports on the clinical severity of the disorder in patients who were homozygous for the haplotype 1 pattern.\(^4\)\(^26\)

In our own study, as in others from northern Europe,\(^10\) haplotype 3 was the only haplotype pattern in strict linkage disequilibrium with a specific mutation (splice donor junction, exon 12). Haplotype 3 is found very rarely on non-PKU chromosomes and so it is likely that the base change arose as a single mutational event and spread throughout northern Europe by population migration and admixture. The low prior probability of a PKU mutation occurring within a rare haplotype is perhaps offset by the fact that a CpG dinucleotide is involved. The splice mutation also occurred on a haplotype pattern which is identical to haplotype 3 except for the loss of an EcoRII recognition site.\(^24\) Although gene conversion or a mutation at this restriction site could be responsible for the appearance of a new haplotype in association with the splice mutation, intragenic crossover 5' to the MspI recognition site is more likely to be the cause. A crossover at this position between a mutant haplotype 3 and a non-mutant haplotype 1, which is common in our population, or a non-mutant haplotype 2, which is less common, would result in the new haplotype pattern.

The R261Q mutation in exon 7, originally described in linkage disequilibrium with haplotype 1 in the Swiss population,\(^15\) was found only on a haplotype 28 allele. Once again a CpG dinucleotide is involved and so it is possible that the mutation arose independently on this pattern. On the other hand, haplotype 28 is very rare at the PAH locus. A single crossover between a haplotype 1 allele carrying the mutation and a haplotype 4 in the region of the gene between...
Molecular heterogeneity at the phenylalanine hydroxylase locus in the population of the south-west of England

The PKU population in the south-west of England is probably representative of the PKU population in England as a whole. Although we have shown that the most common mutations are the splice mutation in exon 12 (26%) and the point mutation in exon 12 (11%), more than 60% of mutant genes do not carry any of the four mutations that were originally found on haplotypes 1, 2, 3, or 4 in other north European populations. To date, at least 20 different mutations have been defined. Some of these are population specific and others are rare mutations that have been found in individual families or communities. It is unlikely, therefore, that another one or two single mutations will account for a large proportion of the remaining 61% of mutant genes in our population. In view of this considerable genetic heterogeneity it is certain that, in the UK at least, molecular genetic techniques will not supersede the biochemical tests which are currently used to identify affected persons in neonatal screening programmes, and large scale screening of sections of the population for carrier status of a mutant PKU gene would be impracticable. Nevertheless, techniques in molecular genetics are the only means by which prenatal diagnosis can be carried out.

We are grateful to Dr E Kunert, Leipzig, Germany, for the positive controls for the haplotype 1 and 4 mutations and to Mr P Rutland, London, for the normal haplotype 4 ASO probe. The work was supported by a grant from the South Western Regional Health Authority and the Shredding Fund. We are indebted to Professor M E Pembrey, London, for helpful discussions in the preparation of the manuscript.

Molecular heterogeneity at the phenylalanine hydroxylase locus in the population of the south-west of England.

L A Tyfield, M J Osborn and J B Holton

*J Med Genet* 1991 28: 244-247
doi: 10.1136/jmg.28.4.244

Updated information and services can be found at:
http://jmg.bmj.com/content/28/4/244

### Email alerting service

*These include:*

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

### Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/