European Journal of Human Genetics (1998) 6, 61–70

ORIGINAL PAPER

Mutation at the phenylalanine hydroxylase gene (*PAH*) and its use to document population genetic variation: the Quebec experience

Kevin C Carter[1,2], Susan Byck[1], Paula J Waters[1,3], Brent Richards[1,2], Piotr M Nowacki[1,2], Rachel Laframboise[5], Marie Lambert[6], Eileen Treacy[1,3,4] and Charles R Scriver[1,2,3,4]

[1]*DeBelle Laboratory, McGill University–Montreal Children's Hospital Research Institute*
Departments of [2]*Biology,* [3]*Human Genetics,* [4]*Pediatrics, McGill University*
[5]*Department of Human Genetics, Centre Hospitalier Université Laval, Ste Foy, Quebec*
[6]*Department of Human Genetics, Hôpital Ste Justine, Montreal*

We describe variation at the *PAH* locus in the population of Quebec. We successfully analyzed 135 of 141 chromosomes from phenylketonuria (PKU) probands (95.7% of the sample), and eight additional chromosomes from a small number of probands with non-PKU hyperphenylalaninemia (HPA). The full set of chromosomes harboured 45 different *PAH* mutations: i) seven polymorphisms (IVS2nt19, IVS3nt-22, IVS6nt-55, Q232Q, V245V, L385L, Y414Y); ii) four mutations causing non-PKU HPA (T92I, E390G, R408Q, D415N); iii) 34 mutations causing PKU. Only six mutations (M1V, R261Q, F299C, S349P, R408W and IVS12nt1) occurred in the whole province at relative frequencies > 5%; most are rare and probably identical by descent. By studying associations of mutations with polymorphic haplotype alleles, we found examples of mutations on different haplotypes that were identical by state, but not by descent because they were recurrent mutations (E280K and R408W); and examples of mutations identical both by state and by descent because of intragenic recombination (S67P, G218V, V245A and IVS12nt1). Ten mutations were first described in Quebec and five are still unique there; three of these 'Quebec' mutations are reported here for the first time (c.125A→T (K42I); [c.470G→A; c.471A→C] (R157N); c.707nt-55 (IVS6nt-55). The *PAH* mutations stratify by geographic region and population, their distributions validating hypotheses about European range expansion to North America during three separate phases of immigration and demographic expansion in the Quebec region over the past four centuries. The *PAH* homozygosity value (*j*) is 0.06 for the total Quebec sample (0.5–0.08 by regions), and the corresponding homoallelic fraction of mutant *PAH* genotypes is 24%. These findings are a documentation of genetic diversity in the Quebec population.

Keywords: mutation; phenylalanine hydroxylase gene (*PAH*); hyperphenylalaninemia; phenylketonuria; population variation

Correspondence: CR Scriver, DeBelle Laboratory, Rm. A-721, McGill University–Montreal Children's Hospital Research Institute, 2300 Tupper Street, Montreal, Quebec H3H 1P3, Canada
Received 11 July 1997; revised 9 September 1997; accepted 15 October 1997

Introduction

Range expansion by Europeans has been a major demographic feature of human history in the current millennium.[1] Three separate phases of European settlement occurred to account for the richly diverse history of the region of North America called Quebec, where resides one of the best documented populations in the world.[2–6] French settlers came before 1759; Loyalists from the United States and emigrants from the United Kingdom came after 1759; non-French, non-British settlers arrived in substantial numbers after 1945. Descendants of these settlers are the majority among the nearly 7 million persons comprising the population of Quebec today, and among them are probands with hyperphenylalaninemia (HPA) due to impaired function of phenylalanine hydroxylase enzyme. The corresponding gene (symbol *PAH*) has been cloned and characterized;[7–10] rare mutations in it cause disease, notably phenylketonuria (PKU),[11] and the disease-causing alleles are widely ascertained by screening newborns for HPA. The probands described in the present report were ascertained in this manner. Accordingly, their allelic variation at the *PAH* locus could be used as markers for the demographic sources of and genetic variation in the population of Quebec.

The *PAH* gene harbours over 328 different mutations (see http://www.mcgill.ca/pahdb); it also has a set of polymorphic alleles in apparent linkage disequilibrium from which polymorphic haplotypes can be generated.[10,12–16] Mutations and polymorphic haplotypes, along with descriptors of attributes and relationships, are recorded in a relational database (*PAH*db) curated and maintained on behalf of the *PAH* Mutation Analysis Consortium.[17] In the present report, the fourth in a series,[18–20] we describe 45 different *PAH* mutations in probands with HPA, some causing PKU, others causing non-PKU HPA, all ascertained by newborn screening in the Quebec population over the past 20 years; they include ten novel mutations first identified in this population. Some mutations, otherwise identical-by-state, provided evidence for recurrent mutation or intragenic recombination to explain their distribution on multiple polymorphic haplotype backgrounds. *PAH* mutations stratify in Quebec by geographic region and population history. Accordingly, our findings have historical and medical as well as genetic interest.

Materials and Methods

The Sample

From a recent epidemiological analysis[21] of newborn screening in Quebec between 1973 and 1990, we ascertained probands with HPA and subdivided them into PKU (classical and variant forms) or non-PKU HPA. The criteria for classification are described elsewhere.[11] We obtained sufficient DNA from 141 independent chromosomes in probands with PKU and completed analysis on 135 (95.7% of the sample). In a related project, we began the analysis of chromosomes from 13 probands with non-PKU HPA; data on eight chromosomes are included here for comparison with the set of PKU-causing mutations. There are no probands of Aboriginal, Oriental or African ancestry in the sample. The project was reviewed and approved for ethics by the Institutional Review Board (Montreal Children's Hospital).

Geographic Features We stratified findings by geographic region. 'Eastern Quebec' is the settled region east of the city of Three Rivers; its population is French Canadian, with a well-documented pattern of settlement and expansion.[22,23] 'Western Quebec' is the corresponding region west of the river; it has a largely French Canadian population and a demographic history different from that in Eastern Quebec. 'Montreal' refers to the region of greater Montreal and environs, where the non-French Canadian component of the Quebec population is predominantly located. The DNA samples were obtained from probands attending the regional medical centers at Laval, Montreal and McGill universities respectively.

Mutation Detection

We used 'broad range', denaturing gradient gel electrophoresis (DGGE) to analyze all exons and flanking regions of intron sequences in the *PAH* gene, after amplification by PCR from designated GC-clamped primers.[24,25] Mutant exons were amplified with non-clamped primers, and sequenced using the dsDNA cycle sequencing system (Gibco BRL). Novel mutations were confirmed on second PCR sample by sequencing, selective restriction analysis, artificially created restriction sites, or hybridization with allele-specific oligonucleotides.

Mutations are named[26,27] using the convention of the *PAH* Mutation Analysis Consortium, i.e. by 'trivial names' which describe the effect of the allele on gene transcription and translation; the corresponding systematic (nucleotide) names are given in the *PAH* Mutation Database.

Haplotype Analysis

We used seven diallelic (RFLP) and two multiallelic polymorphic sites at the *PAH locus*[12,13,15] to describe the relationships with mutations and haplotype background, as described previously.[20]

Calculation of Homozygosity (j)

Homozygosity (j) at the *PAH* locus was calculated for the whole population and for each regional population by the formula $j = \Sigma x_i^2$, where x_i is the frequency of each allele; the uncharacterized alleles (chromosomes) were given the frequency $1/N$ where N is the total number of mutant chromosomes.

The PAH *Mutation Database*

This curated locus-specific mutation database[17,28] is relational in design and contains information on over 328 different mutations (entities), with attributes, relationships (eg to haplotypes and populations), and other details. Data are copied from *PAH*db (Logical Views) and periodically posted on the Internet (http://www.mcgill.ca/pahdb). A search engine is now available (Nowacki P, K Hechtman, CR Scriver – unpublished data). We used the database to compare findings in Quebec and elsewhere.

Terminology

The term 'disease-causing' implies mutations causing hyperphenylalaninemia (HPA). The terms 'chromosome', 'allele' and 'mutation' are used both according to their conventional definitions[29] and in particular contexts here. Chromosome is the operative word describing the physical source in the blood sample of the DNA analyzed at the *PAH* locus (chromosome 12q24.1). Mutation is the equivalent of muton[29] and refers to the smallest unit of DNA changed in the *PAH* nucleotide sequence; each mutation is named according to convention[26,27] (see also http://www.mcgill.ca/pahdb). Allele, as in 'heteroallelic genotype', means the proband inherited two different mutons in *trans* at the *PAH* locus; homoallelic implies identical mutons. A mutation may be identical-by-state on different chromosomes (alleles) but not identical-by-descent; we used mutation-haplotype associations (alleles) to distinguish mutations identical-by-state from those identical-by-descent.

Results

The PAH *Mutations in Quebec*

We identified 45 different *PAH* mutations on 135 PKU chromosomes and on eight additional non-PKU HPA chromosomes in the Quebec probands (Figure 1). Seven mutations, each unlikely to be disease-causing, were found in *cis* with a disease-causing mutation; they are: i) three apparent polymorphisms in introns of which two were previously known (IVS2nt19t→c and IVS3nt-22c→t) and the third, heretofore unknown (IVS6nt-55t→c); ii) four silent polymorphisms in codons (Q232Q, V245V, L385L, Y414Y), all previously known and each useful as an intragenic marker.

The other 38 mutations are apparently disease-causing, many having been analyzed by *in vitro* expression analysis[30] as indicated in Table 1; i) four are

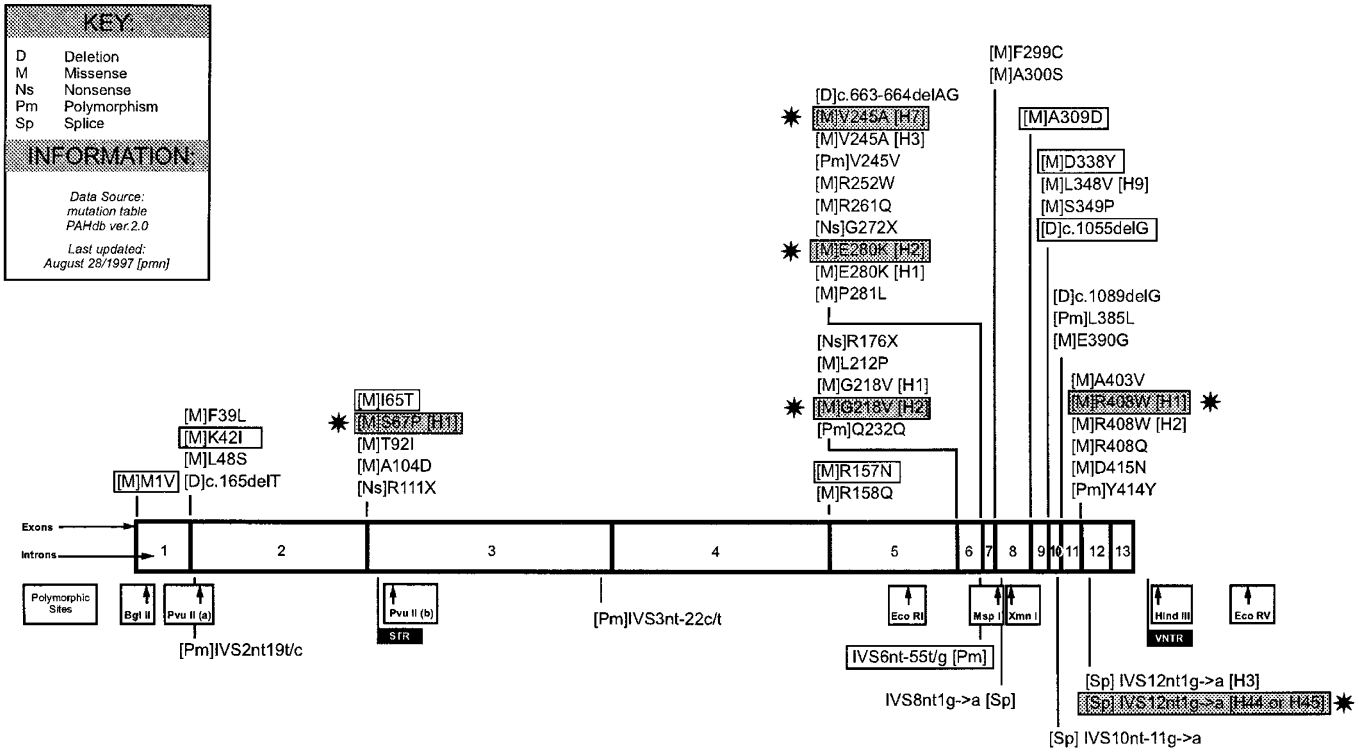

Figure 1 *The human PAH gene (~ 100 kb on chromosome 12q24.1) showing distribution of exons (bars) and introns (white boxes). Below the gene is shown the relative position of diallelic (open boxes) and multiallelic (shaded boxes) polymorphic markers used to construct haplotypes. Forty-five different mutations on Quebec chromosomes are shown. Polymorphic mutations are indicated; the remainder cause hyperphenylalaninemia (HPA): PKU (n = 34), or non-PKU HPA (T92I, E390G, R408G, R408Q, D415N). Mutations that are identical-by-state and found on more than one haplotype ('alleles') are shown in shaded boxes with star. All boxed mutations were first identified in Quebec. Symbols: fs, frameshift; [H], haplotype; numbered as described elsewhere*[53] *© 1997, PAH Mutation Analysis Consortium*

Table 1 PKU-causing *PAH* mutations (and alleles) on chromosomes in Quebec:[a] relative frequencies (%) and distributions by geographic region

Mutation[b] *[H]**[c] *n*[d]	*The Province* *135*	*W. Quebec* *54*	*Montreal* *35*	*E. Quebec* *43*
M1V▲	5.6	1.8	0	14.9
F39L	4.3	8.8	0	2.1
K42I	0.7	0	2.7	0
L48S	2.3	3.5	2.7	0
c.165delT	0.7	1.8	0	0
I65T▲	4.3	3.5	5.4	4.3
S67[H1]*	0.7	0	2.7	0
A104D▲	0.7	0	2.7	0
R111X	0.7	1.8	0	0
R157N▲	0.7	0	0	2.1
R158Q▲	2.8	3.5	2.7	2.1
R176X	0.7	1.8	0	0
L212P	0.7	0	0	2.1
G218V▲[H1]	1.4	0	0	4.3
G218V[H2]*	0.7	1.8	0	0
c.663-664delAG	0.7	0	0	2.1
V245A[H3]	0.7	0	0	4.3
V245A[H7]*	0.7	1.8	0	0
R252W▲	0.7	1.8	0	0
R261Q▲	5.6	1.8	18.9	0
G272X▲	1.4	1.8	0	2.1
E280K▲[H1]	0.7	0	2.7	0
E280K[H2]*	3.5	5.3	0	4.3
P281L▲	2.1	1.8	2.7	2.1
F299C▲	6.4	10.5	5.4	2.1
A300S	0.7	1.8	0	0
IVS8nt1	0.7	0	2.7	0
A309D	2.8	5.3	0	2.1
D338Y	1.4	3.5	0	0
L348V▲	1.4	3.5	0	0
S349P▲	5.7	1.8	2.7	12.8
c.1055delG	0.7	0	2.7	0
c.1089delG	1.4	1.8	2.7	0
IVS10nt-11	1.4	0	2.7	2.1
A403V	0.7	1.8	0	0
R408W▲[H1]*	7.1	5.3	10.8	6.4
R408W[H2]	4.3	1.8	10.8	2.1
IVS12nt▲[H3]	17.1	15.8	10.8	14.9
IVS12nt1[H44,45]*	0.7	0	0	2.1
Sample completed[e] %	95.7	94.7	94.5	91.5
Homozygoisty(*j*)	0.0626	0.0596	0.0852	0.0822

[a]Composite data for all phases of the project (see Methods).
[b]The following 33 mutations, reported in France or French populations[47, 54, 55] identified in the *PAH* mutation database, were not found in the Quebec sample of PKU and non-PKU HPA chromosomes: F55L, IVS2nt5, R68S, P89fs, c.283delATC (delI94), H146Y, I144T, G171A, E178G, Y198fs, G239S, R241C, R241H, R243Q R243X, L249F, R252Q, A259V, C265G, S273F, Y277D, S303P, L311P, P314H, F331L, c.1043del11bp, S350T, c.1066-3c→t(IVS10nt-3) c.1117delGC(A373fs), E390G, c.1199+1g→a(IVS11nt1), c.1220delc(P407fs), R413S, Y414C.
[c]Some mutations were found to be identical-by-state on two different haplotype backgrounds (see text).
*indicates the novel association found in Quebec.
[d]*n* = number of chromosome successfully analyzed with disease-causing mutation identified.
[e]The denominators for calculating percentage of chromosomes successfully analyzed are the original sample sizes: Province, 141; W. Quebec, 57; Montreal, 37; E. Quebec, 47.
▲Mutation has been expressed *in vitro* to confirm an effect on PAH catalytic activity; see[30].

associated with non-PKU HPA of which three (T92I; E390G; D415N) occurred in 'functionally hemizygous' (missense/null) genotypes and the fourth (R408Q) in the homoallelic state; ii) the remaining mutations (n = 34, Table 1) are all associated with PKU and account for 95.7% of the PKU chromosomes analyzed here. Five mutations were identical-by-state on different haplotype backgrounds but not necessarily identical-by-descent (see below). Mutations yet unidentified are not likely to be those undetectable by DGGE, such as large deletions (rare in the *PAH* gene), but they may be located in untranslated regions (5' region or introns) affecting expression in ways yet to be analyzed.

The *PAH* mutation database records 54 disease-causing (PKU) mutations in populations designated as 'France' or 'French'; 33 of these have not yet been found in the Quebec population (Table 1, footnote). Twenty-one of the 'French' *PAH* mutations and many other 'European' alleles were found in Quebec probands.

Mutations and Alleles Particular to Quebec

Ten mutations in the set of 45 in Quebec (Figure 1) were identified first in this population; five were subsequently found elsewhere: M1V in France; I65T in Great Britain, Ireland, Western Europe, Iberia, and in Europeans who emigrated to Australia, and South America; A309D in N. Ireland; c.1055delG in Belgium, Germany and North Africa; R408W on haplotype 1 in Northwest Europe and in Europeans who emigrated to North America and Australia (see *PAH* database). The other five mutations (K42I, R157N, E280K on haplotype 2, D338Y and IVS6nt-55) remain unique to Quebec as of September 1997.

Novel Mutations

Three mutations are reported here for the first time (data seen by reviewers and documented formally elsewhere).[31] Two are disease-causing: c.125A→T (K42I) on haplotype 21; and [c.470G→A; c.471A→C] (R157N) on an undetermined haplotype. Each occurred on a single chromosome. K42I occurred in a proband with classical PKU in a 'functionally hemizygous' *PAH* genotype (K42I/E280K). The R157N mutation, found in a patient with PKU, involves a dinucleotide substitution and it curtails the measured activities of PAH enzyme both *in vitro* and *in vivo*.[32] The third allele is a polymorphism (IVS6nt-55a/c); it was found on three different PKU and control chromosomes (relative frequency 1.4%). Whether it has any effect on phenotype is not known.

Unusual Identical-by-State Mutations

Four PKU-causing mutations were found, in this phase of the project, to be identical-by-state on two different polymorphic haplotype backgrounds (Figure 1, Table 1). Three of the novel relationships (S67P on haplotype 1 *vs* haplotype 4;[33,34] G218V on haplotype 2 *vs* haplotype 1,[24] and V245A on haplotype 7 *vs* haplotype 3[35]) can each be explained by a single intragenic recombination (data shown elsewhere).[31] The fourth mutation (IVS12nt1) was found on either haplotype 44 or 45; this unique association (IVS12nt1 is otherwise almost exclusively on haplotype 3) is unresolved because of technical difficulties with the EcoRV site; nonetheless, a double intragenic recombination would be necessary to explain it.

Two other identical-by-state mutations with unusual haplotype relationships, each first reported in Quebec, were found again in this phase of the project (Table 1). They are R408W on haplotype 1[18,19] and E280K on haplotype 2.[20] In both cases, the finding is compatible with recurrent mutation involving a CpG site.[36, 37]

Mutational Heterogeneity in Quebec

The observed frequency of homoallelic PKU genotypes was 24%. We used the sets of PKU-causing mutations (Table 1) to calculate homozygosity (j) at the *PAH* locus for the Quebec population as a whole and by region. The value for Quebec is 0.06; values are slightly higher in the Montreal and eastern regions and lower in Western Quebec.

Mutation Types in Quebec and the World

The relative frequency distributions of mutation types are similar in Quebec and the world (Figure 2). The great majority of mutations are missense. *PAH* deletions rarely involve more than one or a few nucleotides (http://www.mcgill.ca/pahdb). Insertions are also small and there are none in the Quebec sample. Polymorphisms are not included in Figure 2 because they are underestimated in the world sample whereas their detection has been efficient using DGGE in the Quebec sample.

Effect of Mutations on Phenotype

Whether all 34 mutations associated with PKU (Table 1), or the four mutations associated with non-PKU HPA inherited in functionally hemizygous or homoallelic genotypes (T92I, E390G, R408Q, and D415N, Figure), impair PAH enzyme function *in vivo* is unknown. However, 16 of the PKU mutations and one non-PKU HPA mutation (R408Q) have been studied *in*

vitro by expression analysis[30,32] and Table 1; they all impair catalytic activity. By their nature, the other 20 mutations are also likely to alter PAH enzyme activity. The corresponding *in vivo* genotype-phenotype relationships have been documented in detail elsewhere[38] (see also *PAH* database). Whether all of those designated as neutral polymorphisms are without any phenotypic effect is unknown.

Geographic Stratification of Alleles

There is non-random distribution of *PAH* alleles by geographic region in Quebec (Table 1) and the 12 most prevalent mutations (PKU and non-PKU HPA causing) stratify by geographic region (Figure 3). The distributions are compatible with different demographic histories for the three regions.

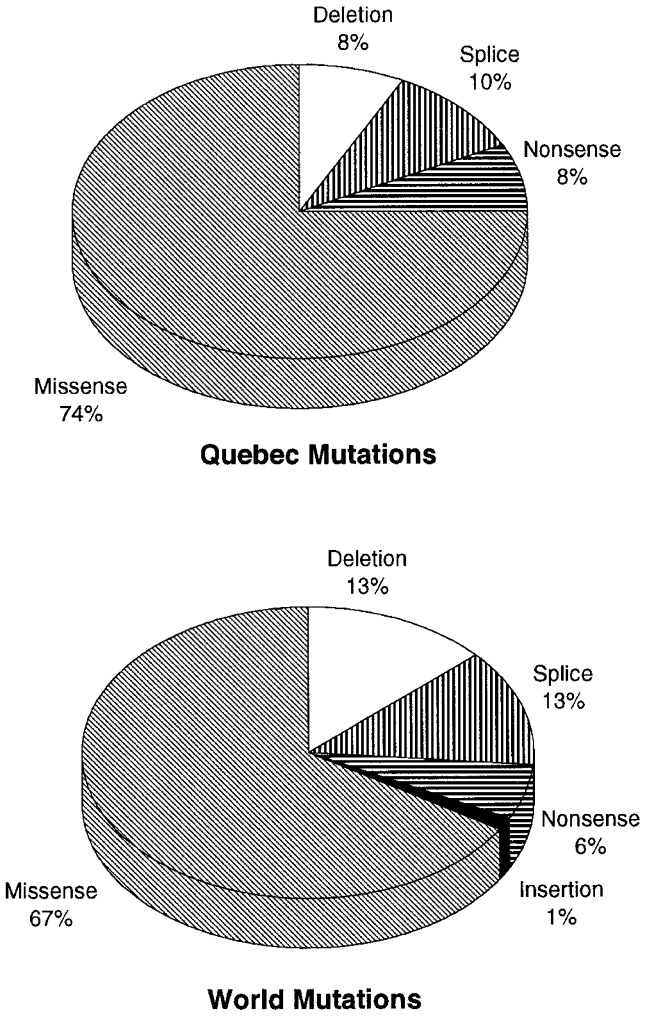

Figure 2 *The Quebec* PAH *mutations excluding polymorphisms, classified by type and compared with the world wide set (*n = *328), the latter taken from the PAH mutation database (http://www.mcgill.ca/pahdb)*

Discussion

Hereditary metabolic diseases, albeit individually rare in prevalence, are important sources of knowledge about human genetic disease in general. From mutation detection by current methods,[39,40] each of the corresponding loci has revealed extensive locus-specific allelic variation. Whereas a few mutations at each locus may be prevalent and shared widely among probands, most disease-causing mutations are rare, their histories coinciding with those of families and defined populations and probably identical by descent.[41] The record of mutations in the *PAH* gene is no exception to this emerging view.

Probands with hyperphenylalaninemia, notably PKU, are being ascertained with high efficiency wherever newborn screening takes place. Accordingly, the sample of mutant chromosomes available for analysis of the *PAH* gene is large, worldwide, originates in the populations of at least 28 countries, and harbours more than 328 different mutations.[17] When these *PAH* mutations are analyzed for their relationships with polymorphic haplotypes, populations of origin, and geographic distributions,[42] the majority that are disease-causing are indeed rare; no more than five being prevalent among Europeans: they are I65T, IVS10nt-11, IVS12nt1, R408W on haplotype 1 and R408W on haplotype 2 (the latter a recurrent mutation)[36] and they account for about two-thirds of the known European *PAH* mutations. Whereas the aggregate frequency of PKU mutations is only about 1% in European genomes,[11,43] individually they are useful markers of human history,[42] notably of European range expansion to Australia,[44] South America[45] and North America.[46] We can now see how their presence in Quebec populations reflects a well-documented range expansion in three different phases after 1600 AD, first from France, then from Great Britain, and lately from other regions of Europe and the world. The Quebec *PAH* mutations (Figure 1) validate demographic and historical hypotheses and interrogate mutation mechanisms.

The rare (disease-causing) and polymorphic (neutral) alleles on PKU chromosomes in the Quebec population (Table 1), tell us the following:

i) The majority (> 80%) of the phenotype-modifying (disease-causing) mutations occur at low

relative frequency ($q < 0.035$) and one third occurs at frequencies below 0.01; among these are the novel mutations (Table 1). No mutation in Quebec occurs at frequencies above 0.18, but five of the 36 PKU mutations account for half of those in the population. These findings are in keeping with emerging views of human genetic variation.[41]

ii) Some *PAH* mutations, known to occur at higher relative frequencies in European populations today, are distributed in all regions of Quebec at corresponding frequencies, eg IVS12nt1 and R408W (Table 1). Other mutations, found in Quebec at the higher frequencies, are non-randomly distributed (stratified by region) (Table 1, Figure 3), eg M1V and S349P in eastern Quebec; R261Q and R408W in Montreal; F39L and F299C in western Quebec. These particular mutations appear to be tracers of historical demography, reflecting the stages of European range expansion to the New World: from France before 1759 (eg M1V), from the British Isles in the second phase after 1759 (R408W on haplotype 1) and from central and southern Europe after 1945 (eg P281L and IVS10nt-11).

iii) The distribution of various *PAH* alleles in French Canadians is compatible with founder effect and genetic drift, for example: a) M1V, a rare allele in

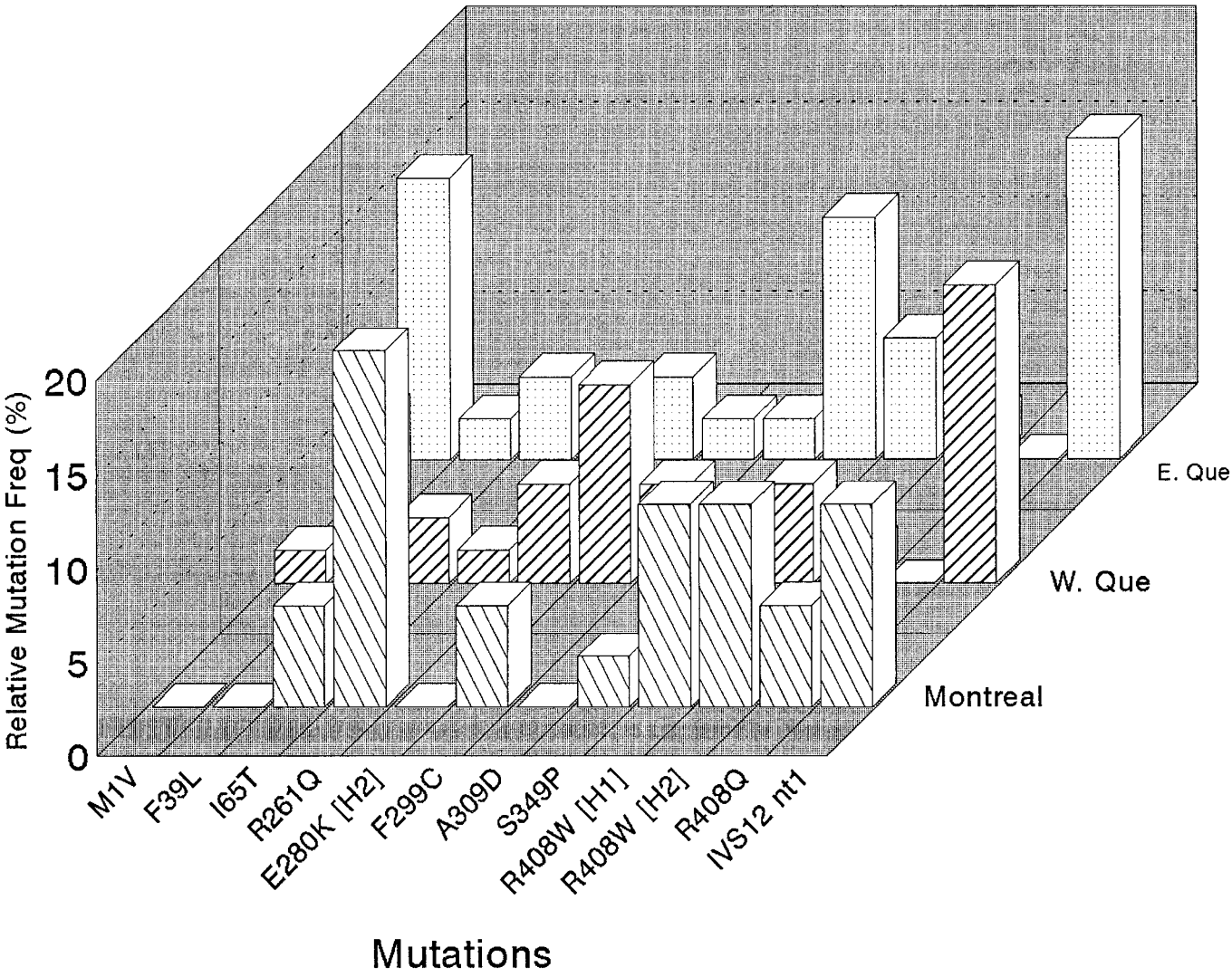

Figure 3 *PAH mutations in PKU probands stratify by geographic region (place of birth) in Quebec. The most prevalent alleles in each region are plotted as relative frequencies in the region; R408Q associates with the non-PKU HPA phenotype, the remainder with classical PKU. Geographic regions are defined in Methods*

France today,[47] and unknown in Europe outside of France, has been traced to ancestors who emigrated from a small region in north-western France[48] who contributed disproportionately to the gene pool in the north-eastern region of New France (Quebec);[49] b) among 54 different PKU mutations found in French populations today, 33 have not been found in Quebec.

iv) The high mutation detection rate on Quebec PKU chromosomes (95.7%) permits us to estimate homozygosity values in this population where the observed fraction (about one quarter) of homoallelic *PKU* genotypes is similar to that found in Europe and elsewhere. The corresponding homozygosity value (*j*) for the Quebec population as a whole, and for each of the regional geographic regions, is low ($j < 0.1$). The Quebec values are similar to those for the United States, the Netherlands and Victoria (Australia) – typical outbred populations – and lower than those for Poland (0.44), Iceland (0.26), Tataria (0.19) and Denmark (0.17) for example;[46] the myth of inbreeding in Quebec appears contradicted here.

v) The Quebec project has generated useful hypotheses about the molecular mechanisms underlying mutations at the *PAH* locus, for example: a) from the discovery of European mutations identical-by-state but on multiple haplotypes in the Quebec population (eg R408W on H1,[18] and E280K on H2)[19] came evidence for recurrent mutations at CpG dinucleotides in the gene.[36, 37] These findings encouraged an analysis of predicted mutability in the coding sequence of the *PAH* gene.[37] b) From the analysis of other mutations again identical-by-state yet on multiple haplotypes has come evidence for intragenic recombination in the 100 kb *PAH* gene; accordingly, these mutations (S67P, G218V, V245A and IVS12nt1) are identical by descent.

vi) Discovery of the recurrent R408W and the I65T mutation in Quebec, along with revealing geneaological reconstructions,[19] led to enquiries about their European origins and how they came to be markers of range expansion. The R408W mutation on haplotype 1, and I65T on haplotype 9, are now seen to cluster in North Western Europe along its Atlantic borders,[50,51] the regions from which settlers of New France/Lower Canada largely came prior to 1945.

vii) The profile of mutation types (Figure 2) at the *PAH* locus in the Quebec poopulation is similar to that found world-wide and in particular to that reported for the United States;[46] family studies have so far ruled out any *de novo PAH* mutation in Quebec. The findings imply that demography and human history, rather than environmental mutagenesis for example, explain the *PAH* mutation profile in these New World populations. Nor is the frequency of unidentified PKU-causing mutations in the Quebec sample exceptional; it may reflect a limitation of the DGGE method, but it is equally likely that these reclusive mutations could occur in the non-coding regions of the gene (although prior searches in the 5' region of the *PAH* gene have been unfruitful on otherwise uninformative 'PKU' chromosomes.[52]

In summary, from a careful study of *PAH* mutations in a well-defined New World population (Quebec), it has been possible to document molecular mechanisms generating genomic diversity (eg recurrent mutation and intragenic recombination) and to trace historical events underlying the genetic variation observed in the population (such as range expansion and genetic drift, during the past 400 years). Furthermore, our model shows how it might be possible, despite the difficulties inherent in extensive allelic heterogeneity for mutation diagnosis,[39,40] to apply knowledge about population genetic variation to efficient prediction, treatment and prevention of a genetic disease.

Acknowledgements

Part of this work was submitted by KCC to the Faculty of Graduate Studies and Research, McGill University in partial fulfilment of requirements for a Master's thesis. It was also presented, in part, at the Annual meeting (29 October–4 November 1996) of the American Society of Human Genetics (Carter K *et al Am J Hum Genet* 1996; **59**: (Suppl.) A30, Abst. p143, 1996). We thank our patients for their contributions to knowledge, Per Guldberg for our introduction to broad-range DGGE, Savio Woo, Randy Eisensmith, and colleagues for their generosity with advice and reagents, Pardeep Kuarah for help with non-PKU HPA patients, Gérard Bouchard, Claude Laberge, Ken Morgan, and Rima Rozen for advice and encouragement during this project. The work was supported in part by the Medical Research Council of Canada (MRC Group in Medical Genetics), the Networks of Centres of Excellence (Canadian Genetics Diseases Network), the Fonds de la Recherche en Santé du Québec (Réseau Génétique Humaine Appliqué), the Interuniversity Institute for Population Research (IREP) and the Ketchum Fund. Lynne Prevost not only helps to curate the *PAH* mutation

database but also controlled the electrons producing this 'manuscript'.

References

1 Crosby AW: *Ecological Imperialism. The Biological Expansion of Europe 900–1900*. Cambridge University Press: Cambridge, 1986.
2 McEvedy C, Jones R: *Atlas of World Population History*; Penguin Books, 1978, 283pp.
3 Charbonneau H, Robert N, Harris RC (eds): *Historical Atlas of Canada*. vol. I: *From the beginning to 1800*. Toronto University Press: Toronto, 1987; The French origins of the Canadian population 1608–1759. Plate 45.
4 Gentilcore RL, Measner D, Doherty D: The coming of the Loyalists. Plate 7. In Gentilcore RL, Measner D, Walder RH (eds). *Historical Atlas of Canada*, vol. II *The land transformed 1800–1891*. Toronto University Press: Toronto, 1993.
5 Cartwright D, MacPherson M: Population Composition. Plate 4. In Kerr D, Holdsworth DS, (eds). *Historical Atlas of Canada*, vol. III, *Addressing the Twentieth Century 1891–1961*. Toronto University Press: 1990.
6 Weaver JC, DeJonge J, Norris D: Trans-Atlantic Migrations. 1831–1851. Plate 9. In Gentilcore RL, Measner D, Walder RH, (eds). *Historical Atlas of Canada*. vol. II, *The land transformed 1800–1891*. Toronto University Press: Toronto; 1993.
7 Woo SLC *et al*: Cloned human phenylalanine hydroxylase gene allows prenatal diagnosis and carrier detection of classical phenylketonuria. *Nature* 1983; **306**: 151–155.
8 Kwok SCM *et al*: Nucleotide sequence of a full-length complementary DNA clone and amino acid sequence of human phenylalanine hydroxylase. *Biochemistry* 1985; **24**: 556–561.
9 Konecki DS *et al*: Structural characterization of the 5' region of the human phenylalanine hydroxylase gene. *Biochemistry* 1992; **31**: 8363–8368.
10 Lidsky A *et al*: Extensive restriction site polymorphism at the human phenylalanine hydroxylase locus and application in prenatal diagnosis of phenylketonuria. *Am J Hum Genet* 1985; **37**: 619–634.
11 Scriver CR, Kaufman S, Eisensmith E, Woo SLC: The hyperphenylalaninemias. In Scriver CR, Beaudet AL, Sly WS, Valle D, (eds). *The Metabolic and Molecular Bases of Inherited Disease*. 7th edn. McGraw Hill Book Co: Place, 1995, pp 1015–1075.
12 Goltsov AA *et al*: Associations between mutations and a VNTR in the human phenylalanine hydroxylase gene. *Am J Hum Genet* 1992; **51**: 627–636.
13 Goltsov AA *et al*: A single polymorphic STR system in the human phenylalanine hydroxylase gene permits rapid prenatal diagnosis and carrier screening for phenylketonuria. *Hum Molec Genet* 1993; **2**: 577–581.
14 Lichter-Konecki U, Schlotter M, Konecki DS: DNA sequence polymorphisms in exonic and intronic regions of the human phenylalanine hydroxylase gene aid in the identification of alleles. *Hum Genet* 1994; **94**: 307–310.
15 Eisensmith RC, Woo SLC: Updated listing of haplotypes at the human phenylalanine hydroxylase (*PAH*) locus. Letter to the Editor. *Am J Hum Genet* 1992; **51**: 1445–1448.
16 Feingold J *et al*: Linkage disequilibrium in the human phenylalanine hydroxylase. *Dev Brain Dysfunct* 1993; **6**: 26–31.
17 Nowacki P, Byck S, Prevost L, Scriver CR: The *PAH* mutation analysis consortium database: update 1996. *Nucl Acids Res* 1997; **25**: 139–142.
18 John SWM *et al*: 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* 1990; **46**: 970–974.
19 Treacy E, Byck S, Clow C, Scriver CR: 'Celtic' phenylketonuria chromosomes found? Evidence in two regions of Quebec province. *Eur J Hum Genet* 1993; **1**: 220–228.
20 Rozen R *et al*: Mutation profiles of phenylketonuria (PKU) in Quebec populations: evidence of stratification and novel mutations. *Amer J Hum Genet* 1994; **55**: 321–326.
21 Lambert DM: The Genetic Epidemiology of Hyperphenylalaninemia in Quebec. Thesis submitted to the Faculty of Graduate Studies and Research, McGill University, Montreal 1994.
22 Bouchard G, DeBraekeleer M: Histoire d'un gènome. Population et génétique dans l'est du Québec. In Bouchard G, DeBraekeleer M (eds): The Sillery, Québec, Presses de l'Univ. du Québec, 1991.
23 Bouchard G, Desjardins-Ouellette M-A, Markowski F, Kouladjian K: La distribution des patronymes au Québec: Témoins des dynamiques de population. *Anthropologie et Sociétés* 1985; **9**: 197–218.
24 Guldberg P, Henriksen KF, Güttler F: Molecular analysis of phenylketonuria in Denmark: 99% of the mutations detected by denaturing gradient gel electrophoresis. *Genomics* 1993; **17**: 141–146.
25 Guldberg P, Güttler F: Mutations in the phenylalanine hydroxylase gene: method for their characterization. *Acta Pediatr Suppl* 1994; **407**: 27–33.
26 Beutler E *et al*: Mutation nomenclature: nicknames, systematic names, and unique identifiers. *Hum Mut* 1996; **8**: 203–206.
27 Ad hoc Committee: Update on nomenclature for human gene mutations. *Hum Mut* 1996; **8**: 197–202.
28 Hoang L, Byck S, Prevost L, Scriver CR. (curators): PAH Mutation Analysis Consortium Database: a database for disease-producing and other allelic variation in the human *PAH* locus. *Nucleic Acids Res* 1996; **24**: 127–231.
29 King RC, Stansfield WD (eds).: *A Dictionary of Genetics*, 4th edn Oxford University Press, New York, 1990.
30 Waters J, Parniak MA, Nowacki P, Scriver CR: In vitro expression analysis of mutations in human and rat phenylalanine hydroxylase: Exploring molecular causes of hyperphenylalaninemia. *Hum Mut* 1997; **11**: 4–17.
31 Carter, KC: Population genetic variation at the human phenylalanine hydroxylase locus. Thesis, McGill University, Faculty of Graduate Studies and Research, Montreal 1996.
32 Waters PJ *et al*: Analysis of the [c.470G- > A; c.471A- > C] (R157N) mutation in human phenylalanine hydroxylase by *in vitro* protein expression in three systems: correlation with *in vivo* phenylalanine oxidation and phenylketonuria phenotype (Abstract). *Am J Hum Genet* 1997.
33 Tyfield LA *et al*: Molecular basis of phenylketonuria in an English population. *Dev Brain Dysfunct* 1993; **6**: 60–67.

34 Dianzani I *et al*: Genetic history of phenylketonuria mutations in Italy. *Am J Hum Genet* 1994; **55**: 851–853.
35 Zschocke J *et al*: Non-phenylketonuria hyperphenylalaninaemia in Northern Ireland: frequent mutation allows screening and early diagnosis. *Hum Mut* 1994; **4**: 114–118.
36 Byck S *et al*: Evidence for origin, by recurrent mutation, of the phenylalanine hydroxylase R408W mutation on two haplotypes in European and Quebec populations. *Hum Molec Genet* 1994; **3**: 1675–1677.
37 Byck S, Tyfield L, Carter K, Scriver CR: Prediction of multiple hypermutable codons in the human *PAH* gene: codon 280 contains recurrent mutations in Quebec and other populations. *Hum Mut* 1997; **9**: 316–321.
38 Kayaalp E: Human *PAH* mutations and hyperphenylalaninemia phenotypes: a meta-analysis of genotype-phenotype correlations. *Am J Hum Genet* 1997; **61**.
39 Cotton RGH: Current methods of mutation detection. *Mut Res* 1993; **285**: 125–144.
40 Grompe M: The rapid detection of unknown mutations in nucleic acids. *Nat Genet* 1993; **5**: 111–117.
41 Weiss KM: Is there a paradigm shift in genetics? Lessons from the study of human diseases. *Molec Phylogenetics and Evolution* 1996; **5**: 259–265.
42 Scriver CR, Byck S, Prevost L, Hoang L: *PAH Mutation Analysis Consortium: variation in the human genome.* John Wiley and Sons: Chichester, 1996; The phenylalanine hydroxylase locus: a marker for the history of phenylketonuria and human genetic diversity. pp.73–96.
43 Penrose LS: Phenylketonuria. A problem in eugenics. *Lancet* 1946; **1**: 949.
44 Ramus SJ, Treacy EP, Cotton RGH: Characterization of phenylalanine hydroxylase alleles in untreated phenylketonuria patients from Victoria, Australia: Origin of alleles and haplotypes. *Am J Hum Genet* 1995; **56**: 1034–1041.
45 Perez B *et al*: Presence of the Mediterranean PKU mutation IVS10 in Latin America. *Hum Molec Genet* 1993; **2**: 1289–1290.
46 Guldberg P, *et al*: Phenylalanine hydroxylase gene mutations in the United States. Report from the Maternal PKU Collaborative Study. *Am J Hum Genet* 1996; **59**: 84–94.
47 Abadie V *et al*: Molecular basis of phenylketonuria in France. *Dev Brain Dysfunct* 1993; **6**: 120–126.
48 Lyonnet S *et al*: Time and space clusters of the French-Canadian M1V phenylketonuria mutation in France. *Am J Hum Genet* 1992; **51**: 191–196.
49 Heyer E, Tremblay M: Variability of the genetic contribution of Quebec population founders associated to some deleterious genes. *Am J Hum Genet* 1995; **56**: 970–978.
50 Eisensmith RC: Recurrence of the R408W mutation in the phenylalanine hydroxylase locus in Europeans. *Am J Hum Genet* 1995; **56**: 278–286.
51 Zschocke J, Mallory JP, Eiken HS, Nevin NC: Phenylketonuria and the peoples of Northern Ireland. *Human Genet* 1997; **100**: 189–194.
52 Svensson E *et al*: Three polymorphisms but no disease-causing mutations in the proximal part of the promotor of the phenylalanine hydroxylase gene. *Eur J Hum Genet* 1993; **1**: 306–313.
53 Eisensmith RC, Woo SLC: Updated listing of haplotypes at the human phenylalanine hydroxylase (*PAH*) locus. *Am J Hum Genet* 1992; **51**: 1445–1448.
54 Bénit P *et al*: Novel frame shift deletions of the phenylalanine hydroxylase gene in phenylketonuria. *Hum Molec Genet* 1994; **3**: 675–676.
55 Bénit P *et al*: Five novel missense mutations of the phenylalanine hydroxylase gene in phenylketonuria. *Hum Mut* 1994; **4**: 229–231.