Globally, \textit{CYP1B1} Mutations in Primary Congenital Glaucoma Are Strongly Structured by Geographic and Haplotypic Backgrounds

Subhabrata Chakrabarti,\textsuperscript{1} Kiranpreet Kaur,\textsuperscript{1} Inderjeet Kaur,\textsuperscript{1} Anil K. Mandal,\textsuperscript{2,5} Rajul S. Parikh,\textsuperscript{3} Ravi Thombis,\textsuperscript{2,5} and Partba P. Majumder\textsuperscript{4}

\textbf{PURPOSE.} To obtain a global perspective on the distribution and evolution of \textit{CYP1B1} mutations in primary congenital glaucoma (PCG) worldwide.

\textbf{METHODS.} Five intragenic single-nucleotide polymorphisms in \textit{CYP1B1}—\textit{R}486G, \textit{A}1198, \textit{V}452L, \textit{D}449D, and \textit{N}453S—were used to generate haplotype data from 138 Indian patients with PCG and 132 ethnically matched normal controls, which were then analyzed in conjunction with data from other populations. Maximum-likelihood estimates of haplotype frequencies were estimated from the genotype data. Subsets of patients and normal control subjects were also genotyped with respect to eight short tandem repeat (STR) markers around the \textit{CYP1B1} locus (\textit{D}2\textit{S}165, \textit{D}2\textit{S}166, \textit{D}2\textit{S}259, \textit{D}2\textit{S}291, \textit{D}2\textit{S}337, \textit{D}2\textit{S}2378, and \textit{D}2\textit{S}2860), to gain evolutionary insights.

\textbf{RESULTS.} Common mutations in \textit{CYP1B1} that are causal of PCG occurred on a uniform haplotype background among Indian patients, which is completely distinct from the modal haplotype background found among unaffected control subjects. Comparison of these data with data from other global regions reveals strong clustering of \textit{CYP1B1} mutations and haplotype backgrounds. The two distinct modal haplotypes found among Indian patients with PCG and control subjects are both ancient with ages of similar magnitudes, as indicated by large variances in the number of repeats at eight STR loci. Together with data from chimpanzee and normal control subjects from India and other global regions, it was possible to make a parsimonious reconstruction of the evolution of these haplotypes.

\textbf{CONCLUSIONS.} The strong association of specific haplotypes with some predominant \textit{CYP1B1} mutations underlying PCG and the observed geographical clustering, probably due to founder effects, may be useful for predictive testing. (\textit{Invest Ophthalmol Vis Sci.} 2006;\textit{47}:43–47) DOI:10.1167/iovs.05-0912

From the \textsuperscript{1}Kallam Anji Reddy Molecular Genetics Laboratory, the \textsuperscript{2}Jasti V. Ramanamma Children’s Eye Care Centre, and the \textsuperscript{3}VST Centre for Glaucoma Care, L. V. Prasad Eye Institute, Hyderabad, India; and the \textsuperscript{4}Human Genetics Unit, Indian Statistical Institute, Kolkata, India.

Supported by Grant BT/PR4774/Med/12/181/2004 from the Department of Biotechnology, Government of India (SC). KK received a predoctoral and IK a postdoctoral fellowship from the Council of Scientific and Industrial Research and the Department of Biotechnology, Government of India.

Submitted for publication July 15, 2005; revised August 30, 2005; accepted November 16, 2005.

Disclosure: S. Chakrabarti, None; K. Kaur, None; I. Kaur, None; A.K. Mandal, None; R.S. Parikh, None; R. Thomas, None; P.P. Majumder, None.

The publication cost of this article was defrayed in part by page charge payment. This article must therefore be marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Corresponding author: Subhabrata Chakrabarti, Brien Holden Eye Research Centre, L. V. Prasad Eye Institute, Road No. 2, Banjara Hills, Hyderabad 500034, India; subhio@lvpei.org.

Primary congenital glaucoma (PCG; online Mendelian Inheritance in Man [OMIM] 231300; http://www.ncbi.nlm.nih.gov/Omim/searchmorbid/, provided in the public domain by the National Center for Biotechnology Information, Bethesda, MD) is a recessively inherited disorder of the eye with variable penetrance. The trabecular meshwork and anterior chamber angle of the eye are affected, leading to the obstruction of aqueous outflow, increased intraocular pressure (IOP) and optic nerve damage. The onset of this disease is seen in the neonatal or early infancy period and, if untreated, results in an irreversible blindness.\textsuperscript{1} The prevalence of PCG varies across populations and continents (1/20,000–1/10,000 in Western populations)\textsuperscript{2} and is higher in inbred populations (1/2500 among the Saudi Arabians,\textsuperscript{3} 1/1250 among the Slovakian Roms,\textsuperscript{2} and 1/3500 among inhabitants of Andhra Pradesh, India)\textsuperscript{3}. Genetic heterogeneity is well documented in PCG and indicated by large variances in the number of repeats at eight STR loci. Together with data from chimpanzee and normal control subjects from India and other global regions, it was possible to make a parsimonious reconstruction of the evolution of these haplotypes.

\textbf{CONCLUSIONS.} The strong association of specific haplotypes with some predominant \textit{CYP1B1} mutations underlying PCG and the observed geographical clustering, probably due to founder effects, may be useful for predictive testing. (\textit{Invest Ophthalmol Vis Sci.} 2006;\textit{47}:43–47) DOI:10.1167/iovs.05-0912

Five intragenic SNPs in \textit{CYP1B1} (GenBank Accession Number U56438—3947 C→G (R486G; exon 2), 4166 G→T (A1198; exon 2), 8131 G→C (V452L; exon 3), 8184 T→C (D449D; exon 3), and 8195 A→G (N453S; exon 3), have been used to generate data on haplotypes associated with the mutant chromosomes in a large set of Slovakian Roms,\textsuperscript{2} Saudi Arabian,\textsuperscript{3} Brazilian,\textsuperscript{11} and American patients with PCG.\textsuperscript{15} All Slovakian Rom patients possessed a single \textit{CYP1B1} mutation (E387K) on a uniform haplotype (G-T-C-C-A) background.\textsuperscript{2} A similar scenario was also noted among Saudi Arabian,\textsuperscript{3} but not among Indonesian,\textsuperscript{9} or Japanese\textsuperscript{12} patients. In the present study, we generated comparable data from India and analyzed these in conjunction with data from other populations to obtain a global perspective on the distribution and evolution of the \textit{CYP1B1}-associated PCG mutations.

\textbf{METHODS}

\textbf{Clinical Details of the Patients and Control Subjects}

The study adhered to the tenets of the Declaration of Helsinki. With approval of the Institutional Review Board, 138 consecutive patients with diagnosed PCG and 132 ethnically matched control subjects, presenting at the L. V. Prasad Eye Institute, Hyderabad, India, from
January 2001 to May 2004, were recruited with informed consent. The diagnosis of PCG was confirmed independently by two surgeons based on clinical examination. In each case of PCG, the eye had an increased corneal diameter (>12.0 mm) with raised IOP (>21 mm Hg) and/or presence of Haab’s striae, or optic disc changes. The ages at onset were at 0 to 1 year of age and symptoms of epiphora and photophobia were the corroborating features. Normal adults without any signs or symptoms of glaucoma and other systemic diseases served as control subjects. Their visual acuity ranged from 20/20 to 20/40, and IOP was <21 mm Hg. Clinical examination on stereo biomicroscopy did not reveal any changes in the optic disc suggestive of glaucoma. The patients and control subjects were matched with respect to their geographical region of habitat.

Molecular Analysis

Peripheral blood samples (5–10 mL) were collected by venipuncture, and DNA was isolated by standard protocols.14 The entire CYP1B1 gene was bidirectionally resequenced on a DNA analyzer (model 3100; Applied Biosystems, Inc. [ABI], Foster City, CA) using appropriate oligonucleotide primers and PCR protocols.15 We have also genotyped subsets of patients and normal control subjects with respect to eight short tandem repeat (STR) loci on chromosome 2, around the CYP1B1 locus, to gain evolutionary insights. The order of the STR loci that were screened is: D2S395, D2S165, D2S367, D2S2259, CYP1B1, D2S391, D2S337, D2S2368, D2S286. (Intermarker genetic distances are provided online in Supplementary Table S1 at http://www.iovs.org/cgi/content/full/47/1/43/DC1.) All the fluorescently labeled STR markers were amplified on a thermal cycler according to the manufacturer’s protocol (model 9700; ABI). The amplicons were pooled according to the corroborating features. Normal adults without any signs or symptoms of epiphora and photophobia were the residual was conserved across different species.

Mutations in CYP1B1 accounted for 44.93% (62/138) of all cases of PCG. The R368H mutation was the most common, accounting for 48.38% of mutations. Further details are provided in Table 1. Seventeen pathogenic mutations in CYP1B1 were observed, of which nine were novel.10,15,19,20 These novel mutations included three frameshift mutations resulting from a 23-bp deletion (g.3905del23bp) and a stop codon due to an insertion of adenine (A) at 30 bp (g.3835) in exon II and a 2-bp deletion (g.7900-7901delCG) in exon III. Five novel missense mutations—A115P, M132R, Q144P, P193L, and S239R—from a 23-bp deletion (g.3905del23bp) and a stop codon due to an insertion of adenine (A) at 30 bp (g.3835) in exon II and a 2-bp deletion (g.7900-7901delCG) in exon III. Five novel missense mutations—A115P, M132R, Q144P, P193L, and S239R—were noted in exon II and one (G466D) in exon III, whose residue is a part of the signature sequence (NH2-FXXGXX-CXG-COOH) and is present in all heme-binding cytochromes. The spectrum of CYP1B1 mutations is larger in the Indian population,10 than among Saudi Arabians,3 Slovakian Roms,2 and Brazilians.11 Marked phenotypic heterogeneity in clinical severity had been observed for these mutations in our earlier studies.10,15,19

Table 2 provides the estimated frequency distributions of haplotypes among patients with PCG, classified by the presence or absence of CYP1B1 mutations, and among unaffected control subjects. These frequency distributions are significantly different ($\chi^2 = 46.84$, 8 df; $P < 0.00001$; infrequent

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>%</th>
<th>SD</th>
<th>%</th>
<th>SD</th>
<th>%</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-G-G-T-A</td>
<td>61.67</td>
<td>6.17</td>
<td>21.86</td>
<td>4.74</td>
<td>17.72</td>
<td>3.52</td>
</tr>
<tr>
<td>C-G-G-C-A</td>
<td>13.33</td>
<td>4.31</td>
<td>29.32</td>
<td>5.22</td>
<td>23.88</td>
<td>3.71</td>
</tr>
<tr>
<td>C-G-G-C-G</td>
<td>4.84</td>
<td>2.72</td>
<td>8.40</td>
<td>2.87</td>
<td>14.32</td>
<td>3.05</td>
</tr>
<tr>
<td>G-T-G-T-A</td>
<td>5.27</td>
<td>2.83</td>
<td>31.86</td>
<td>5.34</td>
<td>4.80</td>
<td>2.56</td>
</tr>
<tr>
<td>G-T-C-C-A</td>
<td>14.89</td>
<td>4.52</td>
<td>1.84</td>
<td>1.54</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>G-T-C-G</td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0.80</td>
<td>2.36</td>
</tr>
<tr>
<td>C-G-G-T-G</td>
<td>0</td>
<td></td>
<td>0.80</td>
<td>2.36</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
haplotypes were pooled to avoid vagaries of small sample sizes). The haplotype frequency distributions among patients with PCG with CYP1B1 mutations (CYP1B1 (+) subgroup) and normal subjects were unimodal, whereas in patients without CYP1B1 mutations (CYP1B1 (−) subgroup) it was bimodal. Further, the modal haplotypes were different for the CYP1B1 (+) subgroup (C-G-G-T-A) and control subjects (G-T-C-C-A). However, the haplotype frequency distributions of the CYP1B1 (−) subgroup and control subjects were strikingly similar (Table 2).

There were many striking differences in frequencies of some haplotypes among the three groups. The C-G-G-T-A haplotype occurred with a high frequency (61.6%) in the CYP1B1 (+) patient subgroup, which is roughly four times higher than that in the control subjects (17.8%; Pexact < 0.00001) and three times higher than in the CYP1B1 (−) patient subgroup (21.8%; Pexact < 0.00001). The difference in frequencies of this haplotype between the CYP1B1 (−) subgroup and control subjects is not statistically significant (Pexact = 0.058). The G-T-C-C-A, which was the modal haplotype (42.3%) among the normal control subjects, showed a decline in frequency (31.9%) in the CYP1B1 (−) subgroup of patients, and a further decline (14.9%) in the CYP1B1 (+) subgroup. These decreases were statistically significant (both Pexact < 0.001). The frequency of the C-G-C-C-A haplotype among normal subjects and the CYP1B1 (−) subgroup is virtually the same, but is significantly lower (Pexact < 0.00001) in the CYP1B1 (+) subgroup.

Table 3 gives the distribution of prevalent CYP1B1 mutations across the four major haplotypes in patients with PCG worldwide. (Complete lists of prevalent and minor CYP1B1-associated PCG mutations on the background of the four major haplotypes are provided in Supplementary Table S2 at http://www.iovs.org/cgi/content/full/47/1/43/DCL.) As is evident from the table, most of the common mutations are clustered on the background of the C-G-G-T-A haplotype. It is striking that specific mutations are generally present on specific haplotype backgrounds, irrespective of geographical location. For example, a major proportion of the R368H mutation that is predominant in India occurs on the C-G-G-T-A background, and it is this same background on which this mutation is also found in diverse geographical areas such as Saudi Arabia3 and Brazil.11 Similarly, the E387K mutation, which is the only mutation present among the Slovakian Roms1 appears on the G-T-C-C-A background and is also found on this same background in United States12 and Brazil.11 Similar features also apply to other mutations, such as 4340delG and R390C. However, it is interesting that although the CYP1B1-associated PCG mutation E387K occurs predominantly on the G-T-C-C-A haplotype in many regions of the world, this haplotype is more strongly associated with the CYP1B1 (−) patient subgroup in India and also occurs in high frequency among normal individuals (Table 2). It may also be noted that unlike in other populations, a relatively minor proportion of R368H, M132R, and E229K mutations occur on multiple haplotype backgrounds among Indian patients with PCG (Table 3). Because it is unlikely that a specific haplotype background favors the recurrence of a specific mutation, this feature is most likely due to founder effects.

**DISCUSSION**

We sought to reconstruct this extreme clustering of CYP1B1 mutations on haplotype backgrounds in patients with PCG worldwide from an evolutionary perspective. The haplotype that is present among the chimpanzee is G-T-G-C-A (http://www.ensembl.org/pan_trogloolades/geneview?gene=ENSPTRG00000011843).
This haplotype has not been found among humans. However, the common haplotype G-T-C-C-A (Table 2) among normal humans is only one mutational step away from the chimpanzee haplotype. This haplotype is obviously an ancient human haplotype. It is unclear whether there is polymorphism in chimpanzees. However, another human haplotype C-G-G-T-A is also common, and the spectrum of CYP1B1 mutations among patients with PCG is also the highest on this haplotype background (Table 3). Although this haplotype is absent among the Roma patients with PCG, it is present in significantly higher frequencies among the patients with PCG of Saudi Arabia, India, and Brazil than in their normal control subjects (Fig. 1). This haplotype can, therefore, be considered an ancient haplotype.

To ascertain whether the ages of the haplotypes C-G-G-T-A and G-T-C-C-A are of similar magnitudes, we screened 20 individuals belonging to each of these two haplotypes with respect to eight STR loci on chromosome 2 (the complete genotype data of the patients and control subjects at these eight loci are presented in Supplementary Table S3 at http://www.iovs.org/cgi/content/full/47/1/43/DC1). The variance in the number of repeats pooled across the eight loci was approximately four in both subsets of individuals, indicating that the ages of these two haplotypes are similar. It is obvious that many of the other haplotypes are recombination derivatives. For example, the two haplotypes C-G-C-C-A (Tables 2, 3) and G-T-G-T-A (Table 3) are best explained as recombination products of the ancestral haplotypes G-T-C-C-A and C-G-G-T-A. Based on considerations of parsimony, we have inferred the evolutionary relationships among the various observed haplotypes associated with the common CYP1B1 mutations in patients with PCG (Fig. 2). Data from other geographical regions will be helpful to refine this scenario.

It is clear from Figure 2, that while the two ancestral haplotypes are found among both patients and control subjects (Table 2), the more ancient PCG-related CYP1B1 mutations occurred on the C-G-G-T-A background, whereas the more recent mutations occurred on the G-T-C-C-A background (since the larger fraction of mutations occur on the C-G-G-T-A background). We, however, note that this inference may be somewhat speculative, but it is not possible to test this speculation until more extensive data become available from wider geographical regions.

The strong association of specific haplotypes with some predominant CYP1B1 mutations underlying PCG and the observed geographical clustering, probably due to founder effects, may be useful for predictive testing. It is unlikely that the causal mutations have arisen multiple times during evolution. The fact that the same mutation is found in multiple geographical regions is probably because of human population movements. For example, the E387K mutation that is highly prevalent among Roma patients with PCG is found in low frequency among patients in the United States and Brazil—a speculation that is bolstered by its association with the same haplotype (G-T-C-C-A) in all three countries. Similarly, it is known that there have been ancient population movements from the Arabian Peninsula to India and immigrants to India have possibly carried the G61E mutation, which is highly frequent in Saudi Arabia. This mutation is found in both countries on the same haplotype (C-G-G-T-A) background. Unfortunately, the reconstruction of the history of ancient population movements is very incomplete to permit a complete explanation of the observed geographical distribution of the PCG-related CYP1B1 mutations based on historical gene flow. Identification of mutations in PCG not due to CYP1B1 and estimates of their frequencies are required from different geographical regions to generate a clearer and broader picture of the haplotypic and geographical clustering of mutations associated with PCG and their evolution.

**Acknowledgments**

The authors thank the patients and their families and the normal volunteers for their participation in the study and Dorairajan Balasubramanian for a critique of the manuscript.

**References**


