Am. J. Hum. Genet. 62:776–784, 1998

Mutation and Polymorphism Analysis of the Human Homogentisate 1,2-Dioxygenase Gene in Alkaptonuria Patients

D. Beltrán-Valero de Bernabé,[1,3] B. Granadino,[1,3] I. Chiarelli,[4] B. Porfirio,[4] E. Mayatepek,[5] R. Aquaron,[6] M. M. Moore,[7] J. J. M. Festen,[8] R. Sanmartí,[9] M. A. Peñalva,[2] and S. Rodríguez de Córdoba[1,3]

[1]Departamento de Inmunología and [2]Departamento de Microbiología Molecular, Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas, and [3]Unidad de Patología Molecular, Fundación Jiménez Díaz, Madrid; [4]Dipartimento di Fisiopatologia Clinica, Unita di Genetica Umana, Universita degli Studi, Firenze; [5]Division of Metabolic Diseases, University Children's Hospital, Heidelberg; [6]Laboratoire de Biologie, Hopital d'Enfants de la Timone, Centre Hospitalier Regional et Universitaire de Marseille, Marseille; [7]Clinical Genetics Service, Institute of Child Health, Bristol Royal Hospital for Sick Children, St. Michael's Hill, Bristol, United Kingdom; [8]Afdeling Reumatologie, Medisch Spectrum Twente, Enschede, The Netherlands; and [9]Servicio de Reumatologia, Hospital Clinic i Provincial de Barcelona, Barcelona

Summary

Alkaptonuria (AKU), a rare hereditary disorder of phenylalanine and tyrosine catabolism, was the first disease to be interpreted as an inborn error of metabolism. AKU patients are deficient for homogentisate 1,2 dioxygenase (HGO); this deficiency causes homogentisic aciduria, ochronosis, and arthritis. We cloned the human *HGO* gene and characterized two loss-of-function mutations, P230S and V300G, in the *HGO* gene in AKU patients. Here we report haplotype and mutational analysis of the *HGO* gene in 29 novel AKU chromosomes. We identified 12 novel mutations: 8 (E42A, W97G, D153G, S189I, I216T, R225H, F227S, and M368V) missense mutations that result in amino acid substitutions at positions conserved in HGO in different species, 1 (F10fs) frameshift mutation, 2 intronic mutations (IVS9-56G→A, IVS9-17G→A), and 1 splice-site mutation (IVS5+1G→T). We also report characterization of five polymorphic sites in *HGO* and describe the haplotypic associations of alleles at these sites in normal and AKU chromosomes. One of these sites, *HGO-3,* is a variable dinucleotide repeat; *IVS2+35T/A, IVS5+25T/C,* and *IVS6+46C/A* are intronic sites at which single nucleotide substitutions (dimorphisms) have been detected; and *c407T/A* is a relatively frequent nucleotide substitution in the coding sequence, exon 4, resulting in an amino acid change (H80Q). These data provide insight into the origin and evolution of the various AKU alleles.

Received December 8, 1997; accepted for publication February 11, 1998; electronically published April 3, 1998.

Address for correspondence and reprints: Dr. S. Rodríguez de Córdoba, Departamento de Immunología, Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas, Velázquez 144, 28006-Madrid, Spain. E-mail: SRdeC@pinar1.csic.es

0002-9297/98/6204-0008$02.00

Introduction

Alkaptonuria (AKU [MIM 203500]), the first disorder to be interpreted as a Mendelian recessive trait (Garrod 1902, 1908), is a rare disease in which homogentisate, an intermediary product in the phenylalanine catabolic pathway, cannot be further metabolized and causes homogentisic aciduria, ochronosis, and arthritis. From early childhood, AKU patients excrete, with their urine, large quantities of homogentisate, which can be spontaneously oxidized to a black pigment that intensely darkens the urine of these patients. The pigment is also deposited in several tissues, causing the black pigmentation (ochronosis) that is characteristic of AKU patients. With time, pigment deposition leads to serious damage to the connective tissues in the spine and large peripheral joints and results in a disabling arthritis, the effects of which appear around the 4th decade of the patient's life and tend to worsen with age (reviewed in La Du 1995). AKU patients are deficient for homogentisate 1,2 dioxygenase(HGO [Enzyme Commission 1.13.11.5]) (La Du et al. 1958), the enzyme that mediates the conversion of homogentisate to maleylacetoacetate (Knox and Edwards 1955). Recently, we cloned the human *HGO* gene and demonstrated that it is the gene for AKU by showing that AKU patients in two Spanish pedigrees were homozygous, or compound heterozygous, for the loss-of-function mutations P230S and V300G (Fernández-Cañón et al. 1996). More recently, two other mutations in the *HGO* gene have been identified in three unrelated Slovak patients, two of whom were reported to be homozygotes for the frameshift mutation 621insG and one for the missense mutation G161R (Gehrig et al. 1997).

The human *HGO* gene maps to chromosome 3 (3q21-23) (Fernández-Cañón et al. 1996). It spans 54,363 bp and codes for a 1,715-nt transcript, which is split in 14 exons ranging from 35 to 360 bp (Granadino et al. 1997). The *HGO* introns, 605–17,687 bp in length, con-

Table 1

PCR Primers for Genomic Amplification of the *HGO* Exons (5′→3′)

Exon	Forward[a]	Reverse	PCR Product (bp)
1	GGGAATTCTGTAACCGGTTAAATAAGCTTG	CTTTAACTCAGCCATTTTCTC	237
2	CATATAGGAATTCTTACCTTAGTAAACTGC	GGCAGATCTTTGGCCTGAAAGCTAGTCATCC	242
3	CTGAATTCGTTGGTGGGAAGGTGGGATGC	CTGTGGATCCCAGGGGGAAGGAGGCAGCAC	249
4	ACTCTAAGGCTTGTATATCTTGTAT	TTCCAATGACCATGGTATTC	230
5	GATGCTATTTATTTAGAAGGTC	GCTTTTGCTTGGCATTCAGGCTGCAAATGG	225
6	GGATTCCAAACGTCCCACCGGTCCC	CAACTATGTATGTGCATATGTGAC	338
7	CAATGATTTGTGTTATGACG	GAAATAAGACAGATTGAAGAATGAAGG	223
8	GGAATTCAAAAGCAAATGCAGCCTTAAGCC	GTTGGATCCAGCAGCAGCTGAAACATCTG	260
9	CGGGCTATGGATCCGGGCAGCTTCTTTATAAC	CTCAAGCGAGGCTTAGAGGCTTG	217
10	GTGGATCCGCTCTTGACATGAAAG	CCTTGCCTGCCAACCCTTTTAC	259
11	CAGAATTCTTACTTCTCCCAAAGG	CTGTGGATCCCTCCCACCCAAGCG	318
12	GCATGAAATGTGTGTCATTGTCC	CTTTGGCTTGCAAATGTGGCTTGG	239
13	GGGGATCCGTGTTTTACTGGTCTTGCCTTG	CATGGTGCCATCGGCAATCCTC	246
14	GGAATTCTTATTGGAAAATTTCACCTACCC	GAAGGGATCCAGGCCAACCACAAGGC	589

[a] In some cases, the *HGO* sequence has been modified to introduce *Eco*RI, *Bam*HI, or *Bgl*II restriction sites (underlined) for subcloning of the amplified product.

tain representatives of the major classes of repetitive elements, including several single-sequence repeats (SSR). Two of these SSRs, a (CT)n repeat in intron 4 (*HGO-1, D3S4496*) and a (CA)n repeat in intron 13 (*HGO-2, D3S4497*), were found to be polymorphic in the Spanish population (Granadino et al. 1997).

Here we have investigated the *HGO* gene in 29 AKU chromosomes from 15 unrelated AKU patients and in 104 normal chromosomes of Spanish origin. We report the detection of 12 previously unreported AKU mutations and the characterization of five novel *HGO* polymorphisms. The analysis of the associations between the AKU mutations and the *HGO* polymorphisms provides useful data for population genetic studies and for investigating the history of the mutations.

Material and Methods

Samples

Fourteen affected individuals were included in these studies. They were diagnosed with AKU on the basis of clinical and radiological examinations and routine analyses for homogentisic aciduria. Four of the patients are from France, three are from Spain, two are from Germany, and one each are from Italy, the United Kingdom, Holland, Algeria, and Turkey. We also studied two obligate carriers, children of an additional AKU patient from France. At least 40 healthy individuals of Spanish origin were screened, to verify that the mutations found in AKU patients were absent in the normal population. This sample size was increased to 52 individuals, to calculate allele frequencies at the *HGO* polymorphic sites and to estimate haplotype frequencies. DNA samples for all AKU patients, their family members, and normal individuals were obtained from peripheral blood lymphocytes, by standard methods. Blood was collected from patients and their relatives after informed consent.

PCR Amplification of HGO *Exons and Sequencing*

Exons of the *HGO* gene were amplified from genomic DNA by use of specific primers derived from the 5′ and 3′ intronic sequences (table 1). The annealing temperature for all primer pairs was 60°C. The corresponding PCR products were purified by agarose gel electrophoresis and extraction with Geneclean II (BIO101). Direct sequencing of PCR products was performed with a dye-terminator cycle-sequencing kit (Perkin-Elmer) using *Taq* FS DNA polymerase. Sequences were resolved on an ABI PRISM 377 automatic sequencer, and the results were analyzed with the ABI Analysis software (version 2.1).

Microsatellite and RFLP Analysis

Analysis of polymorphisms at the *HGO*-1 (*D3S4496*), *HGO*-2 (*D3S4497*), and *HGO-3* (*D3S4556*) microsatellites was performed by PCR, using total human genomic DNA, as described elsewhere. In brief, amplification was performed in a total volume of 20 μl containing 100 ng of genomic DNA; 5 pmol of each primer (see table 3); 1 U *Taq* polymerase (Perkin-Elmer Cetus); 250 μM each dATP, dGTP, and dTTP; 10 μM dCTP; 1 μCi α-[^{32}P] dCTP at 300 Ci mmol^{-1}; 1.0 mM $MgCl_2$ (or 1.25 mM, in the case of *HGO-3*); 50 mM KCl; and 10 mM Tris-HCl (pH 8.3). PCR conditions were two cycles at 94°C for 5 min, 60°C for 1 min, and 72°C for 1 min, followed by 30 cycles (or 25 cycles, in the case of *HGO-3*) at 94°C for 45 s and 60°C for 40 s. Samples were resolved on 6% polyacrylamide sequencing gels (or 8%, in the case of *HGO-3*) and were

Table 2

***HGO* Mutations Identified in AKU Patients**

Mutation	Type	Nucleotide Change[a]	Amino Acid Change/Predicted Consequence	No. of Families	Country or Countries of Origin	Reference
F10 fs	Frameshift	c198GG→ATT	Truncation after Phe10	1	France	Present report
E42A	Missense	c292A→C	Glu42Ala	1	UK	Present report
W97G	Missense	c456T→G	Trp97Gly	1	France	Present report
IVS5+1G→T	Splice site	c509+1G→T	Aberrant splicing	1	Holland	Present report
G152 fs	Frameshift	c621insG	Truncation after Gly152	2	Slovakia	Gehrig et al. (1997)
D153G	Missense	c625A→G	Asp153Gly	1	France	Present report
G161R	Missense	c648G→A	Gly161Arg	1	Slovakia	Gehrig et al. (1997)
S189I	Missense	c733G→T	Ser189Ile	1	Algeria	Present report
I216T	Missense	c814T→C	Ile216Thr	1	Spain	Present report
IVS9-56G→A	Intron change	c817-56G→A	?	2	Italy and France	Present report
IVS9-17G→A	Intron change	c817-17G→A	?	1	Italy	Present report
R225H	Missense	c841G→A	Arg225His	1	Spain	Present report
F227S	Missense	c847T→C	Phe227Ser	1	Spain	Present report
P230S	Missense	c855C→T	Pro230Ser	3	Spain (2) and Turkey	Fernandez-Cañón et al. (1996)
V300G	Missense	c1066T→G	Val300Gly	3	France, Spain, and Germany	Fernandez-Cañón et al. (1996)
M368V	Missense	c1269A→G	Met368Val	2	Germany and France	Present report

[a] Positions of nucleotide changes are from the transcription start site as described in Granadino et al. (1997). The *A*TG initiation codon is located at nucleotide position c168. The complete nucleotide sequence of the human *HGO* transcript has been deposited under accession number AF045167.

exposed on Kodak XAR film, with intensifying screens at −70°C, for 2–12 h.

SSCP Analysis

SSCP analysis was performed by PCR, using total genomic DNA, as described elsewhere (Orita et al. 1989). In brief, amplification was performed in a total volume of 10 μl containing 100 ng of genomic DNA; 12.5 pmol of each primer (see tables 1 and 2); 1 U *Taq* polymerase; 250 μM each dATP, dGTP, and dTTP; 10 μM dCTP; 1 μCi α-[^{32}P]-dCTP at 300 Ci mmol^{-1}; 1.5 mM $MgCl_2$; 50 mM KCl; and 10 mM Tris-HCl (pH 8.3). PCR conditions were one cycle at 94°C for 2 min and 30 cycles of 94°C for 1 min, 60°C for 1 min, and 70°C for 1 min. Samples were resolved on 8% and/or 10% nondenaturing polyacrylamide gels and were exposed on Kodak XAR film, with intensifying screens at −70°C, for 2–10 h.

References for Links among Databases

The gene for alkaptonuria (AKU [MIM 203500]), *HGO,* codes for homogentisate 1,2 dioxygenase. The complete nucleotide sequences for the human *HGO* gene (54,363 bp) and its transcript (1,715 bp) have been deposited under GenBank accession numbers AF000573 and AF045167, respectively.

Results

Mutations in HGO *in AKU Patients*

Fourteen unrelated AKU patients and two obligate carriers, children of another AKU patient, were screened for mutations in the *HGO* gene. This represents a total of 29 novel AKU chromosomes. All 14 exons of the *HGO* gene, including their flanking intronic sequences, were amplified from the genomic DNA of each AKU patient by use of the PCR primers specified in table 1. SSCP analysis was performed on exons 1, 2, 3, 4, 5, 7, 8, and 13, and any exons presenting band shifts were subsequently sequenced. All remaining exons—6, 9, 10 11, 12, and 14—were sequenced in both strands.

Twelve novel *HGO* mutations were identified in 21 of the 29 AKU chromosomes examined. Eight of them were missense mutations, resulting in the following amino acid substitutions: E42A, W97G, D153G, S189I, I216T, R225H, F227S, and M368V; one, F10fs, was a

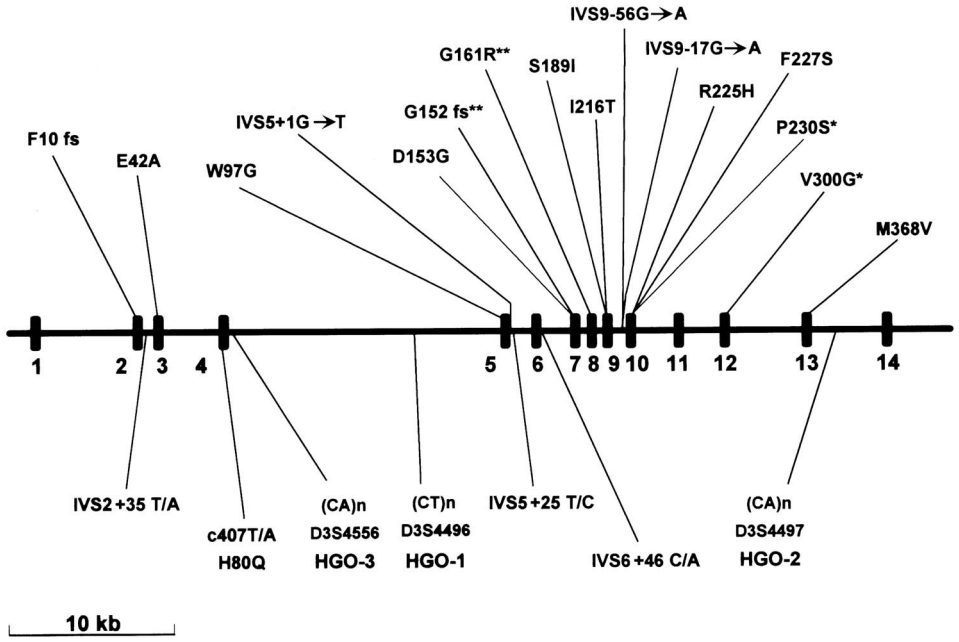

Figure 1 Localization of the AKU mutations and intragenic polymorphic markers at the human *HGO* locus. The human *HGO* gene is shown as a thick horizontal line, with the exons indicated by short, thick vertical lines and labeled 1–14. Above the structure of the *HGO* gene, the thin lines localize the position of each of the 16 AKU mutations characterized thus far. Asterisks indicate mutations that have been described elsewhere (Fernández-Cañón et al. 1996) but have been encountered again in the 29 AKU chromosomes included in the present study. Double asterisks indicate mutations found in a study of AKU patients from Slovakia (Gehrig et al. 1997). Below the structure of the *HGO* gene, the thin lines show the relative positions of seven intragenic polymorphic sites. *HGO-1, HGO-2,* and *HGO-3* are dinucleotide repeats. *IVS2+35T/A, c407T/A, IVS5+25T/C,* and *IVS6+46C/A* are SNPs. *c407T/A* is a relatively frequent nucleotide substitution in the coding sequence (exon 4), causing an amino acid change (H80Q). *HGO-1* (*D3S4496*) and *HGO-2* (*D3S4497*) have been reported elsewhere (Granadino et al. 1997). Distances are drawn to scale.

frameshift mutation (198GG→ATT), two were intronic mutations (IVS9-56G→A, IVS9-17G→A), and one was a splice-site mutation, IVS5+1G→T (table 2 and fig. 1). In addition, five AKU chromosomes harbored loss-of-function mutations that have been described, elsewhere, in two Spanish pedigrees (Fernández-Cañón et al. 1996). Thus, the AKU chromosomes from a German consanguineous pedigree and one AKU chromosome from France carried the V300G mutation. Similarly, the AKU chromosomes in a Turkish consanguineous pedigree had the P230S mutation. In the remaining 3 of the 29 AKU chromosomes (10%), no mutation was detected, however. Two mutations, G161R and G152fs, described, elsewhere, in three Slovak pedigrees (Gehrig et al. 1997) were not found in our group of AKU patients.

None of the 12 novel mutations was observed in a sample of unaffected control individuals, including $\geqslant 80$ chromosomes, and none was coincident in the same haplotype with other *HGO* mutations, which suggests that these changes do not represent frequent polymorphisms. All of the novel missense mutations change amino acid residues that are conserved between humans and other species (i.e., *Aspergillus nidulans* [Fernández-Cañón et al. 1995] and/or *Mus musculus* [Schmidt et al. 1997]). Mendelian inheritance of the mutations was confirmed by SSCP and/or sequence analysis of the relevant amplified DNA fragment in family members (data not shown).

Polymorphic Sites within the Human HGO *Gene*

We have recently described the characterization of two SSRs within the *HGO* gene, a (CT)n repeat in intron 4 (*HGO-1, D3S4496*), and a (CA)n repeat in intron 13 (*HGO-2, D3S4497*) (Granadino et al. 1997). Here we report the identification and analysis of five additional polymorphic sites at the *HGO* locus. These novel polymorphic sites correspond to a variable dinucleotide repeat (*HGO-3*) and four single-nucleotide polymorphisms (SNPs), *IVS2+35T/A, c407T/A, IVS5+25T/C,* and *IVS6+46C/A* (fig. 1). *HGO-3* (*D3S4497*) is a perfect (CA)n repeat located within the 17,687-bp–long intron 4 of *HGO* that was encountered during inspection of the 54-kb–long sequence of the human *HGO* gene (Granadino et al. 1997). The *IVS2+35T/A, c407T/A, IVS5+25T/C,* and *IVS6+46C/A* polymorphisms were

Table 3

Novel Polymorphic Markers within the *HGO* Gene

Marker	Type	HET	PIC	Primer Sequences	Allele	FRQ
HGO-3 (*D3S4556*)	(CA)n	.84	.82	F: GGAGGCAACTTTAGGACAAACTTG	175	.008
				R: CACCCATAAAATTTCCGCTCATCC	177	.008
					187	.085
					189	.208
					191	.192
					193	.200
					195	.100
					197	.123
					199	.062
					201	.015
IVS2 +35A/T	Single base change (intron 2)	.32	.27	F: CATATAGGAATTCTTACCTTAGTAAACTGC	A	.80
				R: GGCAGATCTTTGGCCTGAAAGCTAGT-CATCC	T	.20
IVS5 +25T/C	Single base change (intron 5)	.21	.19	F: GATGCTATTTATTTAGAAGGTC	T	.88
				R: GCTTTTGCTTGGCATTCAGGCTGCAAATGG	C	.12
IVS6 +46C/A	Single base change (intron 6)	.40	.32	F: CATTTTCCTCTGCAATACCTCCATGGAG	C	.72
				R: CAACTATGTATGTGCATATGTGAC	A	.28
c407T/A (H80Q)	Single base change (exon 4)	.30	.25	F: ACTCTAAGGCTTGTATATCTTGTAT	T	.82
				R: TTCCAATGACCATGGTATTC	A	.18

NOTE.—HET indicates heterozygosity; FRQ, frequency; PIC, polymorphism information content.

identified by DNA sequencing while searching for mutations in AKU patients. They were excluded from being AKU mutations because of their elevated frequencies in the normal population. *IVS2+35T/A, IVS5+25T/C,* and *IVS6+46C/A* are SNPs located within *HGO* introns 2, 5, and 6, respectively. The *c407T/A* site is located in exon 4. It corresponds to a relatively frequent T/A variation in the coding sequence, changing a conserved histidine (CAT), at amino acid position 80, for a glutamine (CAA). *c407T/A* is, therefore, an H80Q protein polymorphism. Allele frequencies at the *IVS2+35T/A, c407T/A, IVS5+25T/C, IVS6+46C/A,* and *HGO-3* sites were determined in a sample of 52 normal Spanish individuals and are shown in table 3.

Haplotype Analysis of Polymorphic Sites within the HGO *Gene*

Haplotype associations between alleles at seven polymorphic sites within the *HGO* gene were initially established for a total of 104 chromosomes. This analysis included the *HGO-1* and *HGO-2* sites described elsewhere (Granadino et al. 1997). However, as could be expected, the haplotypes generated from this small sample of chromosomes are almost all unique. The very large number of different *HGO* haplotypes in our sample is mostly due to the very polymorphic microsatellite repeats *HGO-2* and *HGO-3*. Thus, to eliminate excessive variability in the analysis and to obtain useful information about the genetic structure of the Spanish population at the *HGO* locus, we did not consider variation at these two sites.

The analysis of the haplotypic associations of alleles at the *IVS2+35T/A, c407T/A, HGO-1, IVS5+25T/C,* and *IVS6+46C/A* polymorphic sites demonstrated that our population is composed of five major *HGO* haplotypes, which we called A, B, C, D, and E (table 4). The most frequent haplotype is A (~60%), followed by B and C (each ~10%) and then by D and E (each ~7%). As indicated above, each of these haplotypes shows considerable variability at the *HGO-2* and *HGO-3* sites. However, as indicated in table 4, particular *HGO-2* and *HGO-3* alleles are found more frequently associated with each of these haplotypes.

To establish the haplotypic associations between the *HGO* polymorphisms and the AKU mutations, all seven polymorphic markers (*IVS2+35T/A, c407T/A, HGO-3, HGO-1, IVS5+25T/C, IVS6+46C/A,* and *HGO-2*) were characterized in the AKU patients and their families. A complete description of the *HGO* haplotypes for 33 AKU chromosomes, including two pedigrees reported elsewhere (Fernández-Cañón et al. 1996), is presented in table 5.

Discussion

We reported, elsewhere, that the *HGO* gene is the gene for AKU, and we described two loss-of-function mutations, P230S and V300G, in two unrelated Spanish AKU pedigrees (Fernández-Cañón et al. 1996). The P230S mutation was found in three of the four AKU chromosomes included in the original report, which suggests that P230S was a highly prevalent mutation, perhaps characteristic of the Spanish population. The analysis of 29 additional AKU chromosomes—from France (9),

Table 4

Most Representative *HGO* Haplotypes in the Normal Population

Name	Approximate Frequency	*IVS2+35A/T*	*c407A/T*	*HGO-3*[a]	*HGO-1*	*IVS5+25T/C*	*IVS6+46A/C*	*HGO-2*[a]
A	.60	A	T(A)	189 193 195 197	161	T	C	181 187
B	.10	A	T	189 191	161	T	A	179
C	.10	T	T	191	161	T	A	179
D	.07	T	T	191	161	T	C	…[b]
E	.07	A	T	189 193	163	C	A	187

[a] *HGO-2* and *HGO-3* alleles that are found predominantly associated to each of the *HGO* haplotypes.
[b] No specific allele was found associated with this haplotype.

Spain (6), Germany (4), Italy (2), the United Kingdom (2), Holland (2), Algeria (2), and Turkey (2)—demonstrates a large heterogeneity of *HGO* mutations and no prevalence of any single AKU mutation. Thus far, a total of 39 AKU chromosomes, including those in five AKU pedigrees that have been described elsewhere (Fernández-Cañón et al. 1996; Gehrig et al. 1997), have been screened for mutations in *HGO*. Although the number of AKU chromosomes is still small, the count of 16 different mutations identified is notable (table 2). P230S, V300G, and M368V are relatively common mutations that are present in AKU chromosomes from different geographical regions. However, none of these mutations is present in >15% of the AKU chromosomes. Interestingly, only a small percentage (9%, 3 of 33) of the AKU chromosomes analyzed in our laboratory showed no candidate mutation. All of the AKU chromosomes in which no mutation has been identified belong to patients who present, in the other chromosome, one of the *HGO* mutations depicted in table 2, which suggests that these patients are indeed compound heterozygotes for mutations in *HGO* and that there is no genetic heterogeneity in AKU. Experiments are being performed to characterize HGO internal deletions or aberrant splicing in these patients.

In this report, a total of 12 novel mutations are described (table 2). Most of these are missense mutations, changing amino acid residues that are conserved between humans and other species. The effects of these novel mutations on the function of the HGO protein are currently unknown, although experiments are in progress to address this question. Among the novel mutations described here, F10fs is an insertional substitution (198GG→ATT) resulting in a +1 frameshift, after Phe10, that should result in truncation of the HGO polypeptide two residues downstream. Therefore, the F10fs mutation almost certainly represents a null allele. IVS5+1G→T is a splice-site mutation that should alter the correct *HGO* splicing and that probably results in exon skipping and truncation of the HGO polypeptide. The consequences of the IVS9-56G→A and the IVS9-17G→A mutations are, however, more difficult to foresee. The IVS9-56G→A mutation was found in heterozygosis in two unrelated AKU patients. One of them carries the IVS9-17G→A mutation in the other chromosome. Like all other novel mutations described here, the IVS9-56G→A and IVS9-17G→A mutations were not found in a sample of 80 unaffected chromosomes, which indicates that they are not frequent polymorphisms. We cannot rule out the possibility that the IVS9-56G→A and IVS9-17G→A mutations are, however, rare polymorphisms in strong linkage disequilibrium with another not-yet-identified mutation that would be the true causative mutation in these AKU chromosomes. To test whether these mutations prevent the correct mRNA splicing, reverse transcriptase–PCR analysis of the *HGO* mRNA from these patients is being performed.

The mutations thus far characterized in AKU patients are distributed throughout the whole length of the *HGO* gene sequence (fig. 1). It is noticeable, however, that 7 of the 11 missense *HGO* mutations are clustered within exons 7–10 and that 3 of these mutations change amino acid residues at positions 225, 227, and 230, within exon 10. These data point to this short peptide region encoded by exon 10 as crucial for the enzymatic activity of HGO.

The analysis of the haplotypic associations of alleles at the *IVS2+35T/A, c407T/A, HGO-1, IVS5+25T/C,* and *IVS6+46C/A* polymorphic sites demonstrated the presence of five major *HGO* haplotypes in our population sample (table 4). Observed frequencies for the *HGO-A, HGO-B, HGO-C,* and *HGO-D* haplotypes are not different from those expected, if we consider the allele frequencies at the *IVS2+35T/A, c407T/A, HGO-1, IVS5+25T/C,* and *IVS6+46C/A* sites. The *HGO-E*

Table 5

***HGO* Haplotypes Associated to the AKU Alleles**

Mutation	Frequency ($n = 33$)	Associated Polymorphisms								
		IVS2+35A/T	*c407T/A*	*HGO-3*	*HGO-1*	*IVS5+25T/C*	*IVS6+46A/C*	*IVS-9*	*HGO-2*	Origin[a]
P230S	5	A	T	189	163	C	A	817-38insA	185	Spain (1a)
		A	T	189	163	C	A	817-38insA	185	Spain (1b)
		A	A	191	163	C	A	817-38insA	185	Spain (2a)
		T	T	191	163	C	A	817-38insA	185	Turkey (3a)
		T	T	191	163	C	A	817-38insA	185	Turkey (3b)
S189I	2	A	T	189	163	C	A		185	Algeria (4a)
		A	T	189	163	C	A		185	Algeria (4b)
V300G	4	A	T	189	163	C	A		187	France (5a)
		A	T	189	163	C	A		187	Germany (6a)
		A	T	189	163	C	A		187	Germany (6b)
		A	T	189	163	C	A		187	Spain (2b)
M368V	3	A	T	195	161	T	C		183	Germany (7a)
		A	T	195	161	T	C		183	Germany (7b)
		T	T	195	161	T	C		181	France (5b)
E42A	2	A	A	193	161	T	C		187	UK (8a)
		A	A	193	161	T	C		187	UK (8b)
IVS5+1G→T	2	T	T	191	161	T	C		183	Holland (9a)
		T	T	191	161	T	C		183	Holland (9b)
IVS9-56G→A	2	A	T	195	161	T	C		181	Italy (10a)
		A	T	195/199	161	T	C		181	France (11a)
IVS9-17G→A	1	A	T	189	161	T	C		191	Italy (10b)
I216T	1	A	A	193	161	T	C		187	Spain (12a)
D153G	1	A/T	T/A	195/191	161	T	C		187/183	France (13a)
W97G	2	T	T	191	161	T	A		179	France (14a)
		T	T	191	161	T	A		179	France (14b)
R225H	2	T	T	191	161	T	A		183	Spain (15a)
		T	T	191	161	T	A		179	Spain (15b)
F227S	2	A	T	197	161	T	A		181	Spain (16a)
		A	T	197	161	T	A		181	Spain (16b)
F10 fs	1	T	T	191	161	T	A		181	France (17a)
Unknown	3	A	T	199/195	161	T	C		181	France (11b)
		A	T	189	161	T	C		183	France (17b)
		T	T	193	161	T	C		181	Spain (12b)

[a] The chromosomes for each of the AKU patients are indicated in parentheses (i.e., patient 1, chromosomes 1a and 1b).

haplotype, however, appears with a higher frequency than that expected from the individual allelic frequencies. In fact, this haplotype is interesting because it illustrates a situation of strong linkage disequilibrium between the uncommon *HGO-1(163)*, *IVS5+25C,* and *IVS6+46C/A* alleles in our population ($\Delta_{\times 1000} = 66.9$; $P < .001$). Frequencies for these alleles are 0.092, 0.12, and 0.28, respectively.

To test whether these *HGO* haplotypic frequencies were specific to the Spanish population, we also obtained data for the *HGO* intragenic polymorphisms, and we analyzed the haplotypic associations between alleles at four of these sites in an additional population of 46 normal Italian chromosomes. This analysis showed no significant differences between the normal Spanish and Italian samples for the frequencies of the alleles at the *HGO* intragenic markers or their haplotypic associations (data not shown).

There are various possible mechanisms leading to nonrandom associations of alleles that can explain the unexpectedly elevated frequency of the *HGO-E* haplotype in our population. These include selective advantage, random genetic drift, inbreeding, and emigration with amalgamation of population. Analysis of different European and North African populations should help to clarify this point.

The association between AKU mutations and *HGO* haplotypes is very strong, although some variability can be detected at the *HGO-2* and *HGO-3* sites and occasionally at the *IVS2+35T/A* and *c407T/A* sites (table 5). In all four P230S chromosomes that we have identified thus far, a segment of the *HGO* haplotype spanning the *HGO-1, IVS5+25T/C, IVS6+46A/C, IVS9-38insA,* and *HGO-2* loci is fully conserved, with identical alleles in all these sites. This segment includes exon 10, with the P230S mutation (fig. 1 and table 5). These data strongly argue against P230S being a recurrent mutation and support the idea that these four P230S chromosomes are phylogenetically related. Variation at the markers upstream of the *HGO-1* site (*IVS2+35T/A,*

c407T/A, and *HGO-2*) in these AKU chromosomes suggests that P230S is a relatively old mutation. Accordingly, differences among the P230S chromosomes at the *HGO-3* site are probably a reflection of the relatively high rate of mutation at microsatellite sites (10^{-3} per locus per gamete per generation) (Weber and Wong 1993), whereas changes at the *IVS2+35T/A* and *c407T/A* dimorphic sites, with an expected mutation rate of 10^{-9} per site per year, would be better explained by recombination events. A similar explanation is probably true for the variation in the haplotypes associated to the M368V mutation (table 5), which is also interpreted as an old AKU mutation.

The V300G mutation is a relatively common mutation that, like P230S and M368V, has been encountered in more than one geographical region (table 5). Although one of the individuals carrying this mutation belongs to a consanguineous pedigree, it is interesting that the V300G mutation has been found in all instances to be associated with the same *HGO* haplotype, with identical alleles in all seven polymorphic sites. Again, this argues in favor of a common origin for the AKU chromosomes carrying the V300G mutation. The fact that no variation is detected among these AKU chromosomes suggests that V300G originated more recently than P230S.

We have discussed, above, the peculiarities of the *HGO-E* haplotype and have indicated that the strong linkage disequilibrium between the uncommon *HGO-1(163), IVS5+25C,* and *IVS6+46A* alleles can be explained by, among other mechanisms, emigration, with population admixture. It is interesting, in this context, that two relatively common AKU mutations, P230S and V300G, are associated with the *HGO-E* haplotype and that these mutations have been encountered at various geographical locations, including Spain, France, Germany, and Turkey. Furthermore, it is also noticeable that the S189I mutation, encountered in homozygosis in an Algerian pedigree, is also associated with the *HGO-E* haplotype (table 5), perhaps suggesting that the *HGO-E* haplotype is frequent in this North African population.

Several migrations have contributed to the present Spanish population. These include that of the Capsians, who populated northern Africa and southern Europe 20,000 years ago, several migrations from the Sahara around 8,000 years ago, and that of the Arabs, who spread widely throughout the Middle East and northern Africa and who entered the Iberian Peninsula in the 7th century (Ramos-Oliveira 1971). It is tempting to speculate that the *HGO-E* haplotype was introduced in southern Europe by one of these migrations and that the P230S mutation, and perhaps also the V300G mutation, originated in that population prior to its migration into Europe.

The data presented in this article represent the first extensive account of AKU mutations and associated intragenic polymorphisms, and they provide a general understanding of the variability at the *HGO* locus in both AKU and normal populations. This knowledge should be useful for the identification of AKU alleles and for tracing their migration during recent human history.

Acknowledgments

We are grateful to all members of the AKU families for their cooperation and donation of blood samples. We thank Drs. J. Benitez, D. Heine, and A. Díaz, Ms. G. Porras, and Ms. S. Vara de Rey for their help and contribution to this work. We also thank Profs. H. Roux, M. Olmer, and P. C. Acquaviva (Marseille), Dr. T. Mathieu (Nancy), L. Aubert (Toulon), Dr. J. Perterschmitt (Colmar), and Prof. A. Ladjouze (Algier) for identifying AKU patients in France and Algeria. This work was supported by the Fundación Jose Antonio de Castro, the Spanish Comisión Interministerial de Ciencia y Tecnología (grants SAL96/0055 and BIO94/0932), and the Fondo de Investigaciones Sanitarias (grant FIS95/0885). The financial support of the Ministero dell'Università e della Ricerca Scientifica e Tecnologica to B. P. is also acknowledged. In addition, this study is based on work supported by the Fundación Conchita Rabago de Jímenez Díaz under a grant awarded to D. B.-V. de B.

References

Fernández-Cañón JM, Granadino B, Beltrán-Valero de Bernabé D, Renedo M, Fernández-Ruiz E, Peñalva MA, Rodríguez de Córdoba S (1996) The molecular basis of alkaptonuria. Nat Genet 14:19–24

Fernández-Cañón JM, Peñalva MA (1995) Molecular characterization of a gene encoding a homogentisate dioxygenase from *Aspergillus nidulans* and identification of its human and plant homologues. J Biol Chem 270:21199–21205

Garrod AE (1902) The incidence of alkaptonuria: a study in clinical individuality. Lancet 2:1616–1620

———(1908) The Croonian Lectures on Inborn Errors of Metabolism, lecture II: alkaptonuria. Lancet 2:73–79

Gehrig A, Schmidt SR, Muller CR, Srsen S, Srsnova K, Kress W (1997) Molecular defects in alkaptonuria. Cytogenet Cell Genet 76:14–16

Granadino B, Beltrán-Valero de Bernabé D, Fernández-Cañón JM, Peñalva MA, Rodríguez de Córdoba S (1997) The human homogentisate 1,2-dioxygenase (HGO) gene. Genomics 43:115–122

Knox WE, Edwards SW (1955) Homogentisate oxidase of liver. J Biol Chem 216:479–487

La Du BN (1995) Alkaptonuria. In: Scriver CR, Beaudet AL, Sly W, Valle D (eds) The metabolic and molecular bases of inherited disease. McGraw-Hill, New York, pp 1371–1386

La Du BN, Zannoni VG, Laster L, Seegmiller JE (1958) The nature of the defect in tyrosine metabolism in alkaptonuria. J Biol Chem 230:251–260

Orita M, Iwahana H, Kanazawa H, Hayashi K, Sekiya T (1989) Detection of polymorphisms of human DNA by gel

electrophoresis as single-strand conformation polymorphisms. Proc Natl Acad Sci USA 86:2766–2770

Ramos-Oliveira JE (1971) La prehistoria y la antigüedad. In: Historia de España. Vols 1 and 2. Oasis, Mexico City

Schmidt, SR, Gehrig A, Koehler MR, Schmid M, Muller CR, Kress W (1997) Cloning of the homogentisate 1,2-dioxygenase gene, the key enzyme of alkaptonuria in mouse. Mamm Genome 8:168–171

Weber JL, Wong C (1993) Mutation of human short tandem repeats. Hum Mol Genet 2:1123–1128