Am. J. Hum. Genet. 64:1316–1322, 1999

Analysis of Alkaptonuria (AKU) Mutations and Polymorphisms Reveals that the CCC Sequence Motif Is a Mutational Hot Spot in the Homogentisate 1,2 Dioxygenase Gene (*HGO*)

D. Beltrán-Valero de Bernabé,[1,3] F. J. Jimenez,[2] R. Aquaron,[4] and S. Rodríguez de Córdoba[1,3]

[1]Unidad de Patología Molecular and [2]Unidad de Investigación, Fundación Jiménez Díaz, and [3]Departamento de Inmunología, Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas, Madrid; and [4]Laboratoire de Biologie, Hôpital d'Enfants de la Timone, Centre Hôspitalier Regional et Universitaire de Marseille, Marseille, France

Summary

We recently showed that alkaptonuria (AKU) is caused by loss-of-function mutations in the homogentisate 1,2 dioxygenase gene (*HGO*). Herein we describe haplotype and mutational analyses of *HGO* in seven new AKU pedigrees. These analyses identified two novel single-nucleotide polymorphisms (*INV4+31A→G* and *INV11+18A→G*) and six novel AKU mutations (*INV1-1G→A, W60G, Y62C, A122D, P230T,* and *D291E*), which further illustrates the remarkable allelic heterogeneity found in AKU. Reexamination of all 29 mutations and polymorphisms thus far described in *HGO* shows that these nucleotide changes are not randomly distributed; the CCC sequence motif and its inverted complement, GGG, are preferentially mutated. These analyses also demonstrated that the nucleotide substitutions in *HGO* do not involve CpG dinucleotides, which illustrates important differences between *HGO* and other genes for the occurrence of mutation at specific short-sequence motifs. Because the CCC sequence motifs comprise a significant proportion (34.5%) of all mutated bases that have been observed in *HGO,* we conclude that the CCC triplet is a mutational hot spot in *HGO*.

Introduction

Alkaptonuria (AKU; MIM 203500), the first disorder to be interpreted as an inborn error of metabolism (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 (La Du 1995). Patients with AKU are deficient for homogentisate dioxygenase (HGO; EC 1.13.11.5) (La Du et al. 1958), the enzyme mediating the conversion of homogentisate to maleylacetoacetate (Knox and Edwards 1955). We recently cloned the human *HGO* gene and demonstrated that patients with AKU are homozygous or compound heterozygous for loss-of-function mutations in *HGO* (Fernández-Cañón et al. 1996). A total of 17 different AKU mutations have been described thus far (Fernández-Cañón et al. 1996; Gehrig et al. 1997; Beltrán-Valero de Bernabé et al. 1998; Higashino et al. 1998), most of which are missense mutations changing amino acid residues that are conserved between human and other species. Only three mutations (*P230S, V300G,* and *M368V*) have been found in more than one patient.

In addition to the AKU mutations, a number of polymorphisms have been encountered within the human *HGO* gene (Granadino et al. 1997; Beltrán-Valero de Bernabé et al. 1998). The analysis of the haplotypic association of these polymorphic markers in the AKU chromosomes has been very useful for investigating the history of the AKU mutations. These analyses suggested that the most prevalent AKU mutations, *P230S, V300G,* and *M368V,* are not recurrent mutations but instead are probably old mutations that were introduced in Europe with the founder populations and spread throughout western Europe with the different migrations.

Herein we report the results of mutation and polymorphism analyses of the *HGO* gene in seven novel, unrelated AKU pedigrees. These analyses resulted in an identification of novel AKU mutations and *HGO* polymorphisms that prompted us to reexamine all the available *HGO* mutation and polymorphism data, searching for base motifs that could be preferred targets for mutation. We propose, on the basis of these studies, that the CCC motif is a mutational hot spot in *HGO*.

Patients, Material, and Methods

Patients

Seven AKU pedigrees that had not been reported previously were included in our studies. The condition of

Received November 16, 1998; accepted for publication February 19, 1999; electronically published April 12, 1999.

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

© 1999 by The American Society of Human Genetics. All rights reserved.

0002-9297/99/6405-0010$02.00

the probands in these families was diagnosed as AKU on the basis of clinical and radiological examinations and routine analyses for homogentisic aciduria. Three of the pedigrees were from Spain and one each from Poland, the U.S.A., France, and The Netherlands. DNA samples were obtained from peripheral blood lymphocytes of the patients and their family members by use of standard methods. Blood was collected from the patients and their relatives after we obtained informed consent. In one Spanish pedigree, only the children of the proband were available for these studies.

Mutation Screening

The 14 exons of *HGO* were amplified from genomic DNA by use of specific primers derived from the 5′ and 3′ intronic sequences (Beltrán-Valero de Bernabé et al. 1998). SSCP analyses were performed, and any exons presenting band shifts were subsequently sequenced. Briefly, for SSCP analysis, amplification was performed in a total volume of 10 μl containing 100 ng of genomic DNA; 12.5 pmol of each primer; 1 U of *Taq* polymerase (Perkin-Elmer Cetus); 250 μM dATP, dGTP, and dTTP; 10 μM dCTP; 1 μCi of [α-^{32}P]-dCTP at 300 Ci/mmol; 1.5 m*M* $MgCl_2$; 50 m*M* KCl; and 10 m*M* Tris-HCl at pH 8.3. PCR conditions were 1 cycle at 94°C for 2 min, followed by 30 cycles at 94°C for 1 min, 60°C for 1 min, and 72°C for 1 min and 1 cycle at 72°C for 3 min. Samples were resolved on 8% and/or 10% nondenaturing polyacrylamide gels and exposed to Kodak X–Omat AR film with intensifying screens at −70°C for 2–10 h.

Analysis of HGO *Polymorphisms*

We performed PCR analysis of polymorphisms at the *HGO*-1 (*D3S4496*), *HGO*-2 (*D3S4497*), and *HGO*-3 (*D3S4556*) microsatellites by use of total human genomic DNA, as described elsewhere (Beltrán-Valero de Bernabé et al. 1998). The *HGO* intragenic dimorphic markers *IVS2+35A/T, c407T/A, IVS5+25T/C,* and *IVS6+46C/A* were characterized by SSCP analysis, as described elsewhere (Beltrán-Valero de Bernabé et al. 1998).

DNA Sequencing

For each patient, the *HGO* exons selected were amplified from genomic DNA by use of specific primers derived from the 5′ and 3′ intronic sequences, as described elsewhere (Beltrán-Valero de Bernabé et al. 1998). We used the Wizard PCR Preps DNA purification system (Promega) to purify the corresponding PCR products and used a dye terminator cycle sequencing kit (Perkin-Elmer) to perform direct sequencing of them. Sequences were resolved on an ABI PRISM 377 automatic sequencer.

Statistical Methods

To calculate the expected number of mutations associated with the CCC sequence motif, we assumed equal mutability of all nucleotides in the *HGO* sequence. If one of the four nucleotides underlined (NN$\underline{\text{NC}}$C$\underline{\text{CN}}$NN) was mutated at one specific CCC sequence motif, we considered the mutation to be associated to a CCC motif. Therefore, for a given gene sequence, the total number of potential mutations associated to CCC sequence motifs was determined by multiplying by four the number of CCC triplets within the gene sequence. In case of overlapping CCC triplets, the nucleotides that were potential mutation sites for both CCC triplets were counted only once. The expected frequency of mutation at the sites of CCC triplets is calculated as the number of potential mutations at these sites divided by the total number of nucleotides in the sequence. A similar calculation was performed to determine the expected frequency of mutation at CpG dinucleotides. The observed frequency of mutation at CCC and at CpG was contrasted against the expected distribution under the hypothesis of no association (Fleiss 1981). All *P* values reported are two tailed.

Table 1

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

Mutation	Mutation Type	Nucleotide Change	Amino Acid Change/ Predicted Consequence	No. (Origin) of Families	Patient Number (AKU Chromosome)
INV1-1G→A	Splice site	c183-1G→A (exon 2)	Aberrant splicing	1 (Poland)	22 (a)
W60G	Missense	c345T→G (exon 4)	Trp60Gly	1 (USA)	23 (b)
Y62C	Missense	c352A→G (exon 4)	Tyr62Cys	1 (France)	20 (a, b)
A122D	Missense	c532C→A (exon 6)	Ala122Asp	1 (Netherlands)	25 (a)
P230T	Missense	c855C→A (exon 10)	Pro230Thr	1 (Spain)	18 (b)
D291E	Missense	c1040C→A (exon 11)	Asp291Glu	1 (Spain)	21 (a)
M368V[a]	Missense	c1269A→G (exon 13)	Met368Val	4 (2 Spain, 1 USA, and 1 Netherlands)	18 (a), 24 (a, b), 23 (a), and 25 (b)

[a] The *M368V* mutation was described elsewhere in three patients with AKU (Beltrán-Valero de Bernabé et al. 1998).

Table 2

Novel Polymorphic Markers within the *HGO* Gene

Marker	Type	HET	PIC	Primer Sequences, Forward/Reverse	Allele	Frequency
INV4 +31A/G	Single base change (intron 4)	.036	.035	ACTCTAAGGCTTGTATATCTTGTAT/ TTCCAATGACCATGGTATTC	A G	.98 .02
INV11 +18A/G	Single base change (intron 11)	.033	.032	TGGCACGGGAATTATACACCC/ CTGTGGATCCCTCCCACCCAAGCG	A G	.98 .02

References for Links among Databases

The AKU gene, *HGO,* codes for homogentisate 1,2 dioxygenase (*HGO,* EC 1.13.11.5). The complete nucleotide sequences for the human *HGO* gene (54,363 bp) and its transcript (1,715 bp) have been deposited under accession numbers AF000573 and AF045167, respectively. All the information about the coagulation factor VIII gene (*F8C*) was obtained from the haemophilia A database HAMSTeRS.

Results

Novel AKU Mutations

A total of 13 AKU chromosomes in seven AKU pedigrees were screened for mutations in *HGO*. Seven of the 13 AKU chromosomes carry mutations that are described here for the first time. Five of these novel AKU mutations (*W60G, Y62C, A122D, P230T,* and *D291E*) are missense mutations, and one (*IVS1-1G→A*) is a splice-site mutation. With the exception of *W60G,* which was detected only by DNA sequencing, the novel mutations were detected by SSCP and characterized by sequencing. Five of the 13 AKU chromosomes from the seven novel AKU pedigrees carry the *M368V* mutation, a relatively frequent mutation that was encountered previously in three AKU chromosomes from France and Germany (Beltrán-Valero de Bernabé et al. 1998). In only 1 of 13 AKU chromosomes were we not able to identify the mutation, although the 14 *HGO* exons and their intronic flanking sequences were sequenced in both strands. None of the six novel AKU mutations were encountered in a sample of 100 normal chromosomes from the Spanish population, indicating that these nucleotide changes do not represent frequent *HGO* polymorphisms. Table 1 summarizes the mutations encountered in this sample of 13 AKU chromosomes.

Novel HGO *Polymorphisms*

During our search for mutations in the AKU chromosomes, we discovered two single-nucleotide substitutions that are *HGO* polymorphisms. These single-nucleotide polymorphisms are located in intron 4 (*IVS4+31A/G*) and intron 11 (*IVS11+18A/G*) (table 2).

Prevalence of M368V *and Founder Effects*

The results of the analyses of the haplotypic associations of the *HGO* polymorphic markers *IVS2+35A/T, c407T/A, HGO-3, HGO-1, IVS5+25T/C, IVS6+46C/A,* and *HGO-2* in the 13 novel AKU chromosomes included in this report are depicted in table 3. This analysis shows that all five *M368V* AKU mutations are associated with the *HGO-A* haplotype, which was previously found to be associated with the *M368V* mutation in three other AKU chromosomes from France and Germany (Beltrán-Valero de Bernabé et al. 1998). These results strongly suggest that the prevalence of the *M368V* mutation is a consequence of a founder effect, which reinforces the concept that *M368V* is an old mutation that has spread throughout Europe with the different migrations.

Mutational Hot Spots in HGO

Gene mutations are largely related to nucleotide sequences. We have used the MUTPRED program developed by Cooper and Krawczak (1990) to predict the spectrum of point mutations in *HGO*. MUTPRED generates a mutability profile and is based on nearest neighbor–dependent mutation rates estimated from 7,500 single–base pair substitutions from the human gene mutation database.

Analysis of the mutability spectrum of *HGO* generated by MUTPRED showed that 3 (*W60G, W97G,* and *R225H*) of the 16 missense mutations in *HGO* would have been predicted by the program. *R225H* is the only mutation thus far encountered in *HGO* that is associated with a CpG dinucleotide.

Among the *HGO* mutations that were not predicted by MUTPRED, *P230T* is a novel mutation that results from a C→A transversion at *HGO* nucleotide position c855 in exon 10. The *P230T* mutation is interesting because nucleotide position c855 has been shown elsewhere to be involved in a single–base pair change that causes the *P230S* mutation (Fernández-Cañón et al. 1996). Therefore, the same *HGO* nucleotide has been involved in two independent AKU mutations. In one case, the cytosine in position c855 mutates to adenine, generating the *P230T* mutation, whereas in the second case, the same cytosine in *P230S* mutates to thymine to originate the *P230S* mutation. Because it is very unlikely

Table 3

***HGO* Haplotypes Associated to the AKU Mutations**

Mutation	Associated Polymorphisms									Patient[a]
	IVS2+35A/T	*c407T/A*	*IVS4+31A/G*	*HGO-3*	*HGO-1*	*IVS5+25T/C*	*IVS6+46A/C*	*IVS11+18A/G*	*HGO-2*	
INV1-1G→A	A	T	A	201	161	T	C	A	181	PL/22/a
W60G	A/T	T	A	195	161	T	C/A	A	181/179	US/23/b
Y62C	T	T	A	191	161	T	A	A	179	FR/20/a
	T	T	A	191	161	T	A	A	179	FR/20/b
A122D	A	T/A	A	195	161	T	C	G	183	NL/25/a
P230T	A	T	A	199	161	T	C	A	181	SP/18/b
D291E	A	T	A	193/189	161	T	C	A	175	SP/21/a
M368V	A	T	A	199	161	T	C	A	181	SP/18/a
	A	T/A	A	195	161	T	C	A	181	NL/25/b
	A/T	T	A	195	161	T	C/A	A	181/179	US/23/a
	A	T	A	195	161	T	C	A	179	SP/24/a
	A	T	A	195	161	T	C	A	179	SP/24/b
Unknown	A	A	A	195	161	T	C	G	189	PL/22/b

[a] Country of origin/patient number/AKU chromosome. PL = Poland, US = United States, FR = France, NL = The Netherlands, SP = Spain.

that this coincidence occurred by chance, we considered the possibility that position c855 is a mutational hot spot in *HGO* and searched for sequence motifs associated with this position. Position c855 lies within a triplet of cytosines. Interestingly, 34.5% (10/29) of the nucleotide changes (mutations and polymorphisms) thus far characterized in *HGO* lie within, or are adjacent to, a triplet of cytosines or its inverted complement (NNNCCCNNN [the nucleotides that are found mutated at the CCC sequence motifs are underlined]) (table 4). These figures significantly exceed the 9% frequency of mutations at these sites that would be expected on the basis of the frequency of CCC/GGG trinucleotides in *HGO,* suggesting that the CCC motif is a mutational hot spot (table 5).

Discussion

AKU presents a remarkable allelic heterogeneity. Twenty-three different mutations have been identified in a total of 29 AKU pedigrees, which, when one considers that 52% of these patients are homozygous, results in almost one novel mutation per patient. In addition to the 23 AKU mutations, we have identified six single-nucleotide polymorphisms in the *HGO* gene. Therefore, adding mutations and polymorphisms, we have characterized, thus far, 29 different nucleotide changes in the human *HGO* gene.

Gene mutation is a highly nonrandom process. In eukaryotic genomes, ~30% of the point mutations occur in CpG dinucleotides (Cooper et al. 1995). In fact, CpG was first found to be a hot spot for mutation in the factor VIII gene (*F8C*) and subsequently in a wide range of different human genes (Youssoufian et al. 1986). We have analyzed the *HGO* sequence in search of mutational hot spots. Notably, these analyses showed that, whereas a CpG was implicated in only 1 of the 29 nucleotide changes in *HGO,* in 10 of them, the mutated position lies within, or is adjacent to, a triplet of cytosines or its inverted complement. This suggests that the

Table 4

Single-Nucleotide Substitutions at CCC Trinucleotides in *HGO*

Mutation	Mutation Type	Original Sequence (5′→3′) Coding Strand	Mutated Sequence (5′→3′) Coding Strand	Nucleotide Change
D153G	Missense	ATGGGGACT	ATGGGGGCT	c625A→G
G161R	Missense	GAAAGGGAA	GAAAAGGAA	c648G→A
P230S	Missense	GATACCCAT	GATATCCAT	c855C→T
P230T	Missense	GATACCCAT	GATAACCAT	c855C→A
IVS9-56G→A	Intron	CATGGGCAT	CATGGACAT	c817-56G→A
IVS9-17G→A	Intron	TTTGGGTTA	TTTAGGTTA	c817-17G→A
F10fs	Frameshift	TTTGGGAAT	TTTATTGAAT	c198GG→ATT
G152fs	Frameshift	GATGGGGACT	GATGGGGGACT	c621insG
IVS6+46C→A	Polymorphism	AGGGCTTTT[a]	AGGGATTTT	c601+46C→A
IVS11+18A→G	Polymorphism	TGTAGGGGC[a]	TGTGGGGGC	c1046+18A→G

Note.—The nucleotides that are found mutated at the CCC sequence motifs are underlined.
[a] It is unknown which of the two alleles of the polymorphism corresponds to the original sequence.

CCC motif, like the CpG dinucleotide in *F8C,* is a mutational hot spot in *HGO* (table 4).

The increased frequency of mutation at the CCC sequence motif in *HGO* is not a consequence of an abnormally high content in these triplets. The *HGO* gene sequence that we have analyzed, including the sequences of exons and their flanking introns, is 3,609 nt long and contains 34 CpG dinucleotides and 95 CCC trinucleotides. These numbers are similar to those observed in other human genes, such as *F8C* (Youssoufian et al. 1986), in which the CpG dinucleotides are preferred sites of mutation. In *HGO,* the frequency of mutation at CpG dinucleotides does not differ from the random expectation. Instead, it is the frequency of mutation at CCC triplets that is significantly increased, compared with that expected to occur randomly (table 5).

Because most of the single-nucleotide changes that we have encountered in *HGO* are AKU mutations, it could be argued that the increased frequency of mutations observed at CCC (or GGG) motifs could result from a detection bias. To test this possibility, we considered all potential mutations in the *HGO* coding region. This analysis showed that single-nucleotide changes at CCC (or GGG) motifs do not cause missense or nonsense mutations more frequently than do single-nucleotide changes at other, non-CCC motifs. Similarly, we ruled out the possibility that CCC motifs are preferentially associated to evolutionary conserved amino acid positions (identities comparing the human [HGO] and *Aspergillus* [HmgA] amino acid sequences).

Since it is known that the distribution of CpG mutation in the human genome may be nonrandom, the fact that CpG dinucleotides are not mutational hot spots in *HGO* should not be surprising. It has been postulated that this nonrandom distribution of CpG mutational hot spots relates to differences in cytosine methylation in the germ line in different regions of the genome (Reitsma et al. 1993). Whether the low frequency of CpG mutation in *HGO* compared with *F8C* can be attributed to the lower degree of methylation in the CpG dinucleotides in this gene versus those in *F8C* is unknown.

The frequency of nucleotide changes at positions within, or adjacent to, CCC trinucleotides is not significantly increased in other genes with CCC triplet content similar to that of *HGO,* which suggests, as in the case of the CpG dinucleotides, that the CCC putative hot spots may not have a random distribution in the human genome (table 5).

CpG dinucleotides are not the only sequence motif that has been reported to be a mutational hot spot. Among the other examples in which specific short-sequence motifs are associated to mutational hot spots, the most fascinating is the somatic hypermutation of antibody genes (reviewed by Green et al. 1998). In this case, mutations are limited to the DNA in and around the immunoglobulin genes' V regions, within which there are mutational hot spots that have been defined as three- to four-base sequences such as AGC and TAC and their inverted repeats (Smith et al. 1996), with the most prominent motif being RGYW or AGY, where R = purine, Y = pyrimidine, and W = A or T. Although little is known about the molecular mechanisms that are

Table 5

Mutations at CpG and CCC Motifs in *HGO* and *F8C* Genes

Gene	Sequence (bp)[a]	No. at CpG (CpG/kb)	Mutation at CpG (%)[b]		No. at CCC (CCC/kb)	Mutation at CCC (%)[b]	
			Expected	Observed		Expected	Observed
HGO	3,609	34 (9.4)	.5/29 (1.8)	1/29 (3.4)	95 (26.3)	2.6/29 (9)	10/29 (34.5)[c]
F8C	13,191	96 (7.3)	9/678 (1.4) to 5/333 (1.4)	304/678 (44.8)[c] to 65/333 (19.5)	272 (20.6)	48/678 (7.1) to 24/333 (7.1)	87/678 (12.8)[d] to 46/333 (13.8)

[a] The *HGO* sequence included in these analyses corresponds to the 14 exons and the flanking intronic sequences that are amplified by use of the primer pairs described by us elsewhere (Beltrán-Valero de Bernabé et al. 1998). The *F8C* sequence also corresponds to the exons and the flanking intronic sequences and was obtained from GenBank under accession numbers M88628–M88648.

[b] Some mutations in *HGO* have been encountered in more than one AKU patient. However, they were considered only once in these analyses, since the characterization of the *HGO* polymorphisms suggests that their prevalence is a consequence of a founder effect. Nucleotide changes at *F8C* (point mutations, splice mutations, polymorphisms, and deletions/insertions of <3 nucleotides) are described in the *F8C* mutation database (http://europium.mrc.rpms.ac.uk). Because of the difficulty in determining whether the *F8C* mutations that have been encountered more than once are recurrent or identical-by-descent, we considered the two extreme situations: (1) All *F8C* mutations are recurrent, and, therefore, we have included all described cases of *F8C* mutations ($n = 678$); and (2) None of the *F8C* mutations that have been encountered more than once are recurrent mutations, their prevalence being due to founder effects. In this case, only different *F8C* mutations were considered ($n = 333$). Expected frequencies were determined as indicated in Material and Methods.

[c] $P < .001$

[d] Some CCC triplets overlap with CpG dinucleotides in *F8C,* which may cause overestimation of mutations at CCC or CpG.

responsible for this hypermutation process, it is thought that the process may be linked to the elevated transcriptional rate of these genes and that the hypermutation rate may result both from a sequence motif prone to mutate and from some interference with the nucleotide repair machinery (Green et al. 1998). Similarly, in *Escherichia coli,* it has been postulated that the overabundance of mutations induced by alkylating agents at GGG sequence motifs may relate to structural features that make them more refractory to excision or repair mechanisms (Burns et al. 1988). It is unknown whether the overabundance of mutations at CCC motifs in *HGO* is the consequence of a higher mutation rate at these sites or of a deficient repair.

The mutation and polymorphism data for the human *HGO* gene that we have presented here clearly indicate that mutation in CCC triplets occurs in *HGO* at a significantly increased frequency in comparison with that expected to occur randomly. Of all nucleotide changes thus far reported in *HGO,* 34.5% affect CCC trinucleotides, and, notably, one of these sequence motifs has been involved in the generation of two different AKU mutations. Therefore, we propose that the CCC sequence motifs are mutational hot spots in the human *HGO* gene. The data presented also illustrate the differences between genes for the occurrence of mutations at specific short-sequence motifs. These differences may be a consequence of structural (e.g., sequence modification or chromatin organization) and functional (e.g., transcriptional rate) features of the individual genes. An additional series of 13 novel AKU mutations, from Finland and Slovakia, has been reported at various meetings. Analysis of the sequences associated to these additional mutations fully supports our conclusion, stated earlier, that the CCC trinucleotides are mutational hot spots in *HGO*.

Acknowledgments

We are grateful to all members of the families with AKU for their cooperation and donation of blood samples. We thank Drs. M. A. Peñalva, D. Heine, J. Rey-Campos, A. Silva, A. Diaz, Mrs. S. Vara de Rey, and Mrs. L. Gulliksen for their help and their contribution to this work. We also thank Drs. M. Sifry-Platt, M. Jansenss, P. Le François, and R. Le Valliant for identifying patients with AKU in the U.S.A., The Netherlands, France, and Poland. This work was supported by the Fundación Jose Antonio de Castro, the Spanish Comisión Interministerial de Ciencia y Tecnología (SAF96/0055; SAF97/1789E), the Fondo de Investigaciones Sanitarias (FIS98/0687), and the Comunidad Autonoma de Madrid (08.6/0015/1997). In addition, this study is based upon work supported by the Fundación Conchita Rabago de Jímenez Díaz under a grant awarded to D.B.-V. de B.

Electronic-Database Information

Accession numbers and URLs for data in this article are as follows:

GenBank, http://www.ncbi.nlm.nih.gov/Entrez/nucleotide.html (for genomic sequences of *HGO* and its transcript, accession numbers AF000573 and AF045167, respectively, and for the *F8C* sequence, accession numbers M88628–M88648)

HAMSTeRS, http://europium.mrc.rpms.ac.uk (for information about *F8C*)

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim (for AKU [MIM 203500])

References

Beltrán-Valero de Bernabé D, Granadino B, Chiarelli I, Porfirio B, Mayatepek E, Aquaron R, Moore MM, et al (1998) Mutation and polymorphism analysis of the human homogentisate 1, 2-dioxygenase gene in alkaptonuria patients. Am J Hum Genet 62:776–784

Burns PA, Gordon AJ, Kunsmann K, Glickman BW (1988) Influence of neighboring base sequence on the distribution and repair of N-ethyl-N-nitrosourea–induced lesions in *Escherichia coli.* Cancer Res 48(16): 4455–4458

Cooper DN, Krawczak M (1990) The mutational spectrum of single base-pair substitutions causing human genetic disease: patterns and predictions. Hum Genet 85:55–74

Cooper DN, Krawczak M, Antonarakis SE (1995) The nature and mechanisms of human gene mutation. In: Scriver CR, Beaudet AL, Sly W, Valle D (eds) The metabolic and molecular bases of inherited disease. McGraw-Hill, New York, pp 259–290

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

Fleiss J (1981) Statistical methods for rates and proportions. 2d ed. Wiley-Interscience, New York

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

Garrod AE (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

Green NS, Lin MM, Scharff MD (1998) Somatic hypermutation of antibody genes: a hot spot warms up. Bioessays 20(3):227–234

Higashino K, Liu W, Ohkawa T, Yamamoto T, Fukui K, Ohno M, Imanishi H, et al (1998) A novel point mutation associated with alkaptonuria. Clin Genet 53:228–229

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

Reitsma PH, Poort SR, Bernardi F, Gandrille S, Long GL, Sala N, Cooper DN, for the Protein C and S Subcommittee of the Scientific and Standardization Committee of the International Society on Thrombosis and Haemostasis (1993) Protein C deficiency: a database of mutations. Thromb Haemost 69:77–84

Smith DS, Creadon G, Jena PK, Portanova JP, Kotzin BL, Wysocki LJ (1996) Di- and trinucleotide target preferences of somatic mutagenesis in normal and autoreactive B cells. J Immunol 156(7):2642–2652

Youssoufian H, Kazazian HH Jr, Phillips DG, Aronis S, Tsiftis G, Brown VA, Antonarakis SE (1986) Recurrent mutations in haemophilia A give evidence for CpG mutation hot spots. Nature 324:380–382