HUMAN MUTATION 22:43–50 (2003)

RESEARCH ARTICLE

Independent Mutational Events Are Rare in the ATM Gene: Haplotype Prescreening Enhances Mutation Detection Rate

Midori Mitui,[1] Catarina Campbell,[1] Gabriela Coutinho,[1,2] Xia Sun,[1] Chih-Hung Lai,[1] Yvonne Thorstenson,[3] Sergi Castellvi-Bel,[1] Luis Fernandez,[1] Eugenia Monros,[4] Beatriz Tavares Costa Carvalho,[5] Oscar Porras,[6] Gumersindo Fontan,[7] and Richard A. Gatti[1]*

[1]Department of Pathology and Laboratory Medicine, The David Geffen School of Medicine, Los Angeles, California; [2]Instituto de Biofisica Carlos Chagas Filho, UFRJ, RJ, Brazil; [3]Stanford Genome Technology Center, Stanford University Medical School, Palo Alto, California; [4]Hospital Sant Joan de Deu, Barcelona, Spain; [5]Department of Pediatrics, UNIFESP-Escola Paulista de Medicina, SP, Brazil; [6]Immunology Service, National Children Hospital Dr Carlos Saenz Herrera, San Jose, Costa Rica; [7]Immunology Unit Hospital La Paz, Madrid, Spain

Communicated by Jacques S. Beckmann

Mutations in the *ATM* gene are responsible for the autosomal recessive disorder ataxia-telangiectasia (A-T). Many different mutations have been identified using various techniques, with detection efficiencies ranging from 57 to 85%. In this study, we employed short tandem repeat (STR) haplotypes to enhance mutation identification in 55 unrelated A-T families of Iberian origin (20 Spanish, 17 Brazilian, and 18 Hispanic-American); we were able to identify 95% of the expected mutations. Allelic sizes were standardized based on a reference sample (CEPH 1347-2). Subsequent mutation screening was performed by PTT, SSCP, and DHPLC, and abnormal regions were sequenced. Many STR haplotypes were found within each population and six haplotypes were observed across several of these populations. Single nucleotide polymorphism (SNP) haplotypes further suggested that most of these common mutations are ancestrally related, and not hot spots. However, two mutations (8977C>T and 8264_8268delATAAG) may indeed be recurring mutational events. Common haplotypes were present in 13 of 20 Spanish A-T families (65%), in 11 of 17 Brazilian A-T families (65%), and, in contrast, in only eight of 18 Hispanic-American families (44%). Three mutations were identified that would be missed by conventional screening strategies. In all, 62 different mutations (28 not previously reported) were identified and their associated haplotypes defined, thereby establishing a new database for Iberian A-T families, and extending the spectrum of worldwide ATM mutations. Hum Mutat 22:43–50, 2003. **© 2003 Wiley-Liss, Inc.**

KEY WORDS: ataxia-telangiectasia; ATM; Iberian; haplotype; masked mutations; haplotype prescreening; hot spots; mutation analysis

DATABASES:
ATM – OMIM: 208900; GenBank: U82828; www.benaroyaresearch.org/bri_investigators/atm.htm (ATM mutation database)

INTRODUCTION

Ataxia-telangiectasia (A-T; MIM# 208900) is characterized by progressive cerebellar ataxia, ocular apraxia, conjunctival telangiectasias, immunodeficiency, recurring sinopulmonary infections, chromosomal instability, cancer predisposition, and radiation hypersensitivity [Gatti et al., 2001]. A-T is transmitted as an autosomal recessive, affecting 1 in 40,000–100,000 children. Heterozygotes are at an increased risk for malignancy, particularly breast cancer [Swift et al., 1991; Easton, 1994; Athma et al., 1996; Dork et al., 2001; Chenevix-Trench et al., 2002; Spring et al., 2002; Concannon, 2002; Sommer et al., 2002]. The frequency of A-T heterozygotes is estimated at 1% of the population [Swift et al., 1991; Taylor et al., 1994], although recent reports of additional missense type ATM mutations in cancer patients suggest that this figure may be significantly higher [Gatti et al., 1999; Concannon, 2002].

Identifying mutations in the ATM gene has been difficult due to its large size [Concannon and Gatti, 1997]. The gene extends over 150 kb of genomic DNA,

Received 3 December 2002; accepted revised manuscript 4 March 2003.

*Correspondence to: Richard A. Gatti, MD, The David Geffen School of Medicine, Department of Pathology, 675 Young Drive South, Los Angeles, CA 90095-1732. E-mail: rgatti@mednet.ucla.edu

Grant sponsor: Ataxia-Telangiectasia Medical Research Foundation; Grant sponsor: NIH; Grant number: NS36323.

DOI 10.1002/humu.10232
Published online in Wiley InterScience (www.interscience.wiley.com).

includes 66 exons, and has an open reading frame of 9168 nt. The *ATM* gene product contains 3,056 amino acids and is a member of the phosphatidylinositol (PI) 3-kinase family of proteins, with the kinase domain in its C-terminal region [Savitsky et al., 1995]. The *ATM* gene plays a key role in several pathways involved in cell-cycle control, oxidative stress, and DNA repair [Gatti et al., 2001; Shiloh and Kastan, 2001].

Mutations in A-T patients are present throughout the *ATM* gene and lead mainly to null mutations with loss of ATM protein. Founder effect haplotypes and ATM mutations have been described for some populations [Gilad et al., 1996; Stankovic et al., 1998; Telatar et al., 1998a; Telatar et al., 1998b; Laake et al., 1998]. In this study, 55 A-T patients of Iberian origin were characterized from Spain, Brazil, and a Hispanic-American population.

Haplotypes were first defined using four short tandem repeat (STR) markers (D11S1819, NS22, D11S2179, and D11S1818) spanning a region of ~1.4 cM. Markers NS22 and D11S2179 are located within the ATM gene. A total of 105 of the 110 expected mutations (95%) were identified within these three populations, including 62 different mutations. This strategy allowed us to extend the mutation spectrum and to demonstrate the efficiency of haplotype prescreening in identifying mutations. Haplotype prescreening also identified some masked mutations that would not have been detected by any conventional screening method.

MATERIALS AND METHODS

The reference sequence for ATM used was GenBank U82828. Primers and PCR conditions may be obtained from the authors on request.

Patients

Fifty-five unrelated A-T patients of Iberian origin were studied, consisting of 20 individuals from Spain, 17 from Brazil, and 18 from a Hispanic-American population, primarily from Mexico. All patients had classical A-T phenotypes and laboratory confirmation of the diagnosis. Blood samples were collected according to approved human subject protection protocols.

Haplotype Analysis

STR genotyping of patients and their families was performed using four microsatellite markers from chromosome 11q22-23: D11S1819, NS22, D11S2179, and D11S1818 [Rotman et al., 1994; Uhrhammer et al., 1995; Vanagaite et al., 1995; Udar et al., 1999]. In each case, forward primer was end-labeled with γ^{33}P-ATP. PCR products were run on 6% denaturing polyacrylamide gel, which was then dried and exposed to x-ray film. Allelic sizes were standardized to a reference sample (CEPH 1347-02), for which the absolute sizes had been pre-determined by direct sequencing: S1819 (137/137), NS22 (163/165), S2179 (139/147), and S1818 (160/162). SNP genotyping was performed by single strand conformation polymorphism (SSCP) at three polymorphic sites (IVS17–56G>A, 5557G>A, and IVS62–55T>C) that together define the major SNP haplotypes in the ATM region: H2 (GGT), H3 (GAT), and H4 (AGC) [Thorstenson et al., 2001; Campbell et al., 2003].

Mutation Screening

When RNA was available, samples were screened for mutations by protein truncation testing (PTT) [Telatar et al., 1996], followed by SSCP [Castellvi-Bel et al., 1999] on genomic DNA. Samples in which mutations were not detected after screening by the above methods were subjected to denaturing high performance liquid chromatography (DHPLC) analysis [Thorstenson et al., 2001]; PCR fragments included ~90 nucleotides of adjacent intronic sequences. Messenger RNA was isolated from EBV-transformed lymphoblastoid cell lines (LCLs) using RNeasy kit (Qiagen, Valencia, CA). PCR products were purified by QIAquick PCR purification kit (Qiagen) and sequenced using the Thermo Sequenase™ Radiolabeled Terminator Cycle Sequencing Kit (USB, Cleveland, Ohio) or automated sequencing. Nucleotide numbering is based on +1 being the first translational start codon.

RESULTS

Mutations were detected for 105 of 110 alleles (95%). The five mutations that were missed by PTT and SSCP were still missed by DHPLC. Sixty-nine different haplotypes were identified in the 55 families studied. Phase was determined by haplotyping the parents of the patient or by comparing the haplotypes of patients with a shared mutation. In four cases the phase could not be defined by either method. Haplotypes IBERIAN [1]–[6] recurred across the three different populations and were also compared to the haplotypes of 41 Costa Rican families.

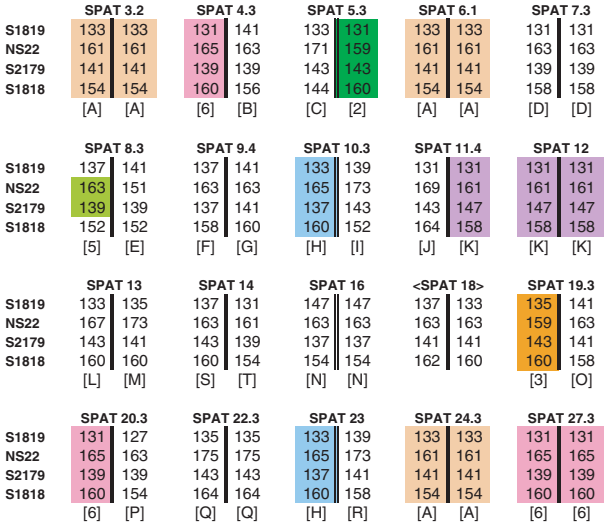
Spanish A-T (SPAT) Families (N=20)

Twenty-six different haplotypes were identified in 40 affected chromosomes (Fig. 1A). Of seven recurring haplotypes (color shaded), three were specific for Spain (SPAT [A], [H], and [K]), and four were also observed in other Iberian populations (IBERIAN [2], [3], [5], and [6]). Sixty-five percent of Spanish families carried at least one of these recurring haplotypes. In general, specific mutations were found on specific haplotypes. Mutations were identified for 38 of the 40 alleles (95%). Of 22 different mutations (Fig. 1B) there were 14 truncations, five aberrant splicings, one missense, one replacement of the stop codon with a phenylalanine plus three additional amino acids, and one genomic deletion of 3450 nt (described below).

Brazilian A-T (BRAT) Families (N=17)

Twenty-one different haplotypes were identified in 34 affected chromosomes (Fig. 2A). Of five recurring haplotypes, one was specific for Brazil (BRAT [I]) and four were also observed in other Iberian populations (IBERIAN [1], [2], [3], and [5]). Sixty-five percent of Brazilian A-T families carried at least one recurring haplotype. Mutations were identified for 32 of 34 alleles (94%). Of 19 different mutations there were 14 truncations, three aberrant splicings, one genomic deletion of 3450 nt, and one genomic deletion of ~17 kb.

A heterozygous patient, AT162LA, carried two mutations located in the same region of the gene. Both were missed by DHPLC. One allele consisted of a genomic deletion of 3450 nt, that excluded exons 29, 30,

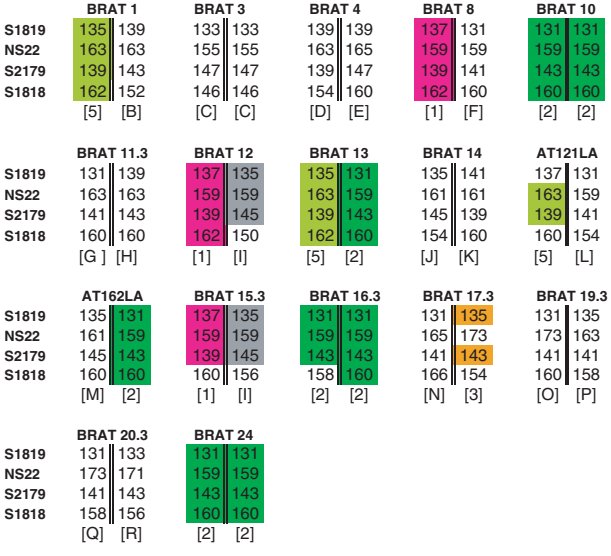


(A)

Ex/Int	Patient	SPAT Haplo	Mutation	Consequence
8	SPAT 10.3[a]	I	640delT	Termination
9	SPAT 7.3[h]	D	**715delT**	Termination
16	SPAT 13[a]	M	2250G>A	Aberrant splicing
18	SPAT 10.3[b]	H	2413C>T	Termination
18	SPAT 23[a]	H	2413C>T	Termination
IVS21	SPAT 8.3[a]	E*	IVS21+1G>A	Aberrant splicing
IVS21	SPAT 9.4[a]	G*	IVS21+1G>A	Aberrant splicing
27	SPAT 5.3[a]	C	3712_3716delTTATT	Termination
28	SPAT 23[b]	R	**3836G>A**	Termination
28	SPAT 22.3[h]	Q	3894_3895insT	Termination
IVS28	SPAT 5.3[b]	2	**IVS28+1711del3450**	Exons 29~31 deleted
IVS33	SPAT 20.3[a]	P	IVS33+2T>C	Aberrant splicing
37	SPAT 19.3[a]	O	5188C>T	Termination
39	SPAT 19.3[b]	3	**5644C>T**	Termination
45	SPAT 14	T	**6342_6343 insC + 6347_6348delGG**	Aberrant splicing
57	SPAT 13[b]	L	**8100A>T**	Lys>Asn (K2700N)
57	SPAT 9.4[b]	F	**8103_8104delAA**	Termination
58	SPAT 4.3[a]	B	8177C>T	Termination
58	SPAT 8.3[b]	5	8264_8268delATAAG	Aberrant splicing
59	SPAT 16[h]	N	8283_8284delTC	Termination
64	SPAT 18		**8875_8878delGACT**	Termination
64	SPAT 4.3[b]	6*	8977C>T	Termination
64	SPAT 20.3[b]	6	8977C>T	Termination
64	SPAT 27.3[h]	6	8977C>T	Termination
64	SPAT 11.4[a]	K*	8977C>T	Termination
64	SPAT 12[h]	K	8977C>T	Termination
65	SPAT 3.2[h]	A	9010_9037del28	Termination
65	SPAT 6.1[h]	A	9010_9037del28	Termination
65	SPAT 24.3[h]	A	9010_9037del28	Termination
65	SPAT 11.4[b]	J	**9170_9171delGA**	Stop codon>Phe+3aa

(B)

FIGURE 1. **Haplotypes and mutations of 20 Spanish families. A: Recurring haplotypes are color shaded, and were assigned numbers instead of letters whenever they were also observed in non-Spanish families. The < > indicate that phase has not been defined. B: Spanish mutations corresponding to affected haplotypes (A). Bolded mutations have not been reported previously. Underlined amino acids are conserved. Asterisks indicate where the same mutation occurred on different haplotypes. Superscripts: a, first allele; b, second allele; h, homozygote. Nucleotide numbering is based on +1 being the A of the first translational start codon.**



(A)

Ex/Int	Patient	BRAT Haplo	Mutation	Consequence
11	BRAT14[a]	J	**1110C>G**	Termination
12	BRAT4[a]	E	1563_1564delAG	Termination
26	AT121LA[a]	L	**3485T>G**	Termination
28	BRAT8[a]	1	3802delG	Termination
28	BRAT12[a]	1	3802delG	Termination
28	BRAT15.3[a]	1	3802delG	Termination
IVS28	BRAT4[b]	D	IVS28+1G>A	Aberrant splicing
IVS28	BRAT10[h]	2	**IVS28+1711del3450**	Exons 29 ~ 31 deleted
IVS28	BRAT13[a]	2	**IVS28+1711del3450**	Exons 29 ~ 31 deleted
IVS28	AT162LA[a]	2	**IVS28+1711del3450**	Exons 29 ~ 31 deleted
IVS28	BRAT16.3[h]	2	**IVS28+1711del3450**	Exons 29 ~ 31 deleted
IVS28	BRAT24[h]	2	**IVS28+1711del3450**	Exons 29 ~ 31 deleted
29	AT162LA[b]	M	**4002_4005delCTTA**	Termination
31	BRAT19.3[a]	O	**4303A>T**	Termination
33	BRAT20.3[a]	R	**4732C>T**	Termination
39	BRAT17.3[a]	3	**5644C>T**	Termination
53	BRAT11	H	7517_7520delGAGA	Termination
IVS54	BRAT14[b]	K	IVS54-3T>G	Aberrant splicing
55	BRAT20.3[b]	Q	7792C>T	Termination
55	BRAT8[b]	F	7884_7888del5	Termination
55	BRAT12[b]	I	**7913G>A**	Termination
55	BRAT15.3[b]	I	**7913G>A**	Termination
58	BRAT1	5	8264_8268delATAAG	Aberrant splicing
58	BRAT13[b]	5	8264_8268delATAAG	Aberrant splicing
58	AT121LA[b]	5	8264_8268delATAAG	Aberrant splicing
59	BRAT17.3[b]	N	8395_8404del10	Termination
61	BRAT19.3[b]	Z	**8620C>T**	Termination
IVS63	BRAT3[h]	C	IVS63del17kb	Exons 64 ~ 65 deleted

(B)

FIGURE 2. **Haplotypes and mutations of 17 Brazilian families. A: Recurring haplotypes are color shaded, and were assigned numbers instead of letters whenever they were also observed in non-Brazilian families. B: Brazilian mutations corresponding to affected haplotypes (A). Bolded mutations have not been reported previously. Superscripts: a, first allele; b, second allele; h, homozygote. Nucleotide numbering is based on +1 being the A of the first translational start codon.**

and the first 34 nt of exon 31 (Fig. 3A, 3B); the second allele was a truncating mutation in exon 29 (4002–4005delCTTA). The genomic deletion of 3450 nt in the first allele created a homozygous pattern for the second allele when detected by SSCP and sequencing, because of the absence of PCR primer annealing sites flanking exons 29–31 on the deleted allele. The deletion was first appreciated in homozygous Brazilian patients (BRAT 10,

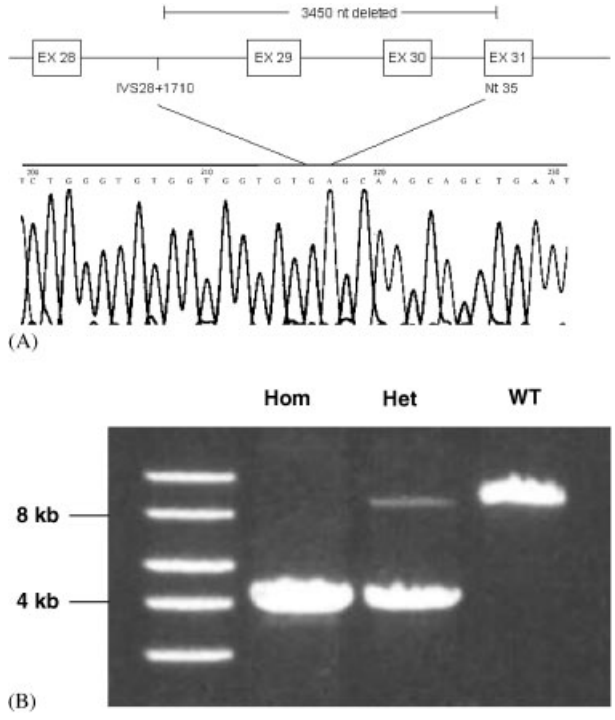


FIGURE 3. **Mutation associated with Iberian Haplotype [2] detected on SPAT 5.3, BRAT 10, BRAT 13, AT162LA, BRAT 16, and BRAT 24. A: Absence of 3450 nt observed at the genomic level. The large genomic deletion of 3450 nt starts at IVS28+1711 (GenBank U82828: 73863) and excludes nucleotides 1–34 of exon 31 (77312). B: PCR of genomic DNA performed with primers located in the regions flanking the deletion. The wild-type fragment of ~8 kb and a smaller sized fragment of ~4.5 kb are present in heterozygous patients carrying the deletion. Homozygous patients show only the smaller sized fragment.**

16.3, and 24) with the haplotype IBERIAN [2] (Fig. 2A).

Another mutation consisted of a ~17-kb genomic deletion that begins within the LINE-1 repeat in intron 63 and ends in a LINE-1 repeat beyond the 3′UTR of the gene. We reported a similar mutation in Costa Rican A-T families (CRAT [B]) [Telatar et al., 1998a].

Hispanic-American A-T (HAAT) Families (N=18)

Twenty-eight different haplotypes were identified in 36 affected chromosomes (Fig. 4A). Of seven recurring haplotypes, three were specific for the Hispanic-American group (HAAT [H], [J], and [M]), and four were also observed in other Iberian populations (IBERIAN [1], [3], [4], and [6]). In contrast to the Spanish and Brazilian families, recurring haplotypes were observed in only eight of 18 Hispanic-American families (44%). Haplotype [H], found in two HAAT patients (AT42LA and AT196LA), carried the 103 C>T truncating mutation that was identified previously as a founder effect among North African Jews [Gilad et al., 1996; Campbell et al., 2003]. Haplotype IBERIAN [4] carrying the null mutation 5908C>T, had been identified previously as the most common Costa Rican founder haplotype, CRAT [A] [Telatar et al., 1998a]. In two Hispanic-American families (AT10LA and AT195LA), a deletion of four intronic nucleotides (IVS20-579del-AAGT) resulted in the inclusion of 65 nt of intronic sequence in the mRNA, identical to a pseudo-exon mutation described recently by Pagani et al. [2002], in a German patient.

Mutations were identified for 35 of 36 alleles (97%). Of 27 different mutations (Fig. 4B), there were 16 truncations, eight aberrant splicings, two missenses, and one inclusion of a pseudo-exon.

Costa Rican A-T (CRAT) Families

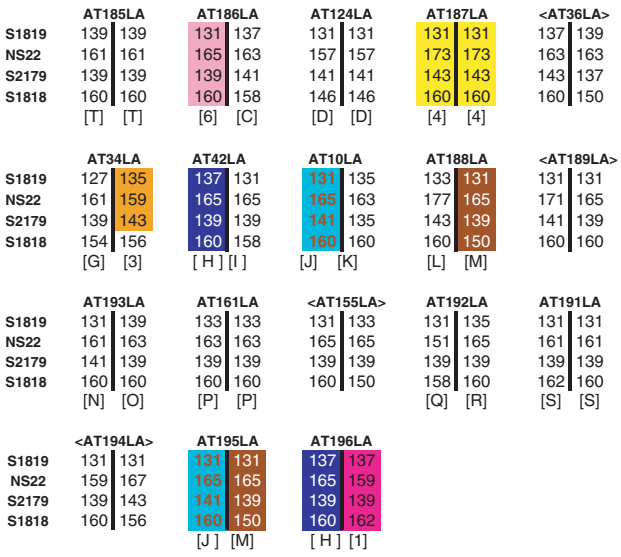
In previously published studies of 82 affected chromosomes, only 10 different haplotypes were defined [Uhrhammer et al., 1995; Telatar et al., 1998a], making this the most homogeneous of the Iberian A-T populations studied to date. Four founder haplotypes accounted for 88% of the affected chromosomes in that country. Table 1 shows the six most common founder haplotypes, newly defined by standardized allelic sizes, and their associated mutations. Because the CRAT [A] haplotype was also observed in a Hispanic-American family, it has been redesignated as IBERIAN haplotype [4], as discussed above. Mutations have been defined for 93% of the 82 chromosomes. Of six different mutations there were three truncations, two aberrant splicings, and one genomic deletion of ~17 kb [Telatar et al., 1998a].

Identical Mutations on Identical STR and SNP Haplotypes

Four STR haplotypes (IBERIAN [1]–[4]), and the corresponding ATM mutations, were observed repeatedly among A-T patients of Spanish, Brazilian, Hispanic-American, and Costa Rican backgrounds. SNP haplotyping provided only limited information [Bonnen et al., 2000; Thorstenson et al., 2001]. Patients with a shared mutation and STR haplotype, in general, also shared a common SNP haplotype (e.g., all four patients with the IBERIAN [4] STR haplotype and mutation 5908C>T, also shared the H4 SNP haplotype, Table 2A). Taken together, these data suggest a common ancient ancestry for each of the Iberian mutations.

Identical Mutations on Different STR Haplotypes

We observed four exceptions to the general rule of finding identical mutations on identical haplotypes: 1) IVS21+1G>A on haplotypes SPAT [E] and SPAT [G]; 2) 8264_8268delATAAG on haplotypes IBERIAN [5] and CRAT [E]; 3) 8977C>T on haplotypes IBERIAN [6] and SPAT [K]; and 4) IVS63del~17kb on haplotypes BRAT [C] and CRAT [B] (Table 2B). Two of these mutations were observed on different SNP haplotypes as well: 8264_8268delATAAG on SNP haplotypes H2 (four patients) and H4 (one patient), and 8977C>T on SNP haplotypes H2 (four patients) and H3 (two patients) (Table 2B). Since these mutations



(A)

Ex/Int	Patient	HAAT Haplo	Mutation	Consequence
5	AT42LA[a]	H	103C>T	Termination
5	AT196LA[a]	H	103C>T	Termination
9	AT193LA[a]	O	790delT	Termination
IVS9	AT155LA[a]		**IVS9+1G>A**	Aberrant splicing
11	AT36LA[a]		1110delC	Termination
11	AT188LA[a]	L	**1235G>C**	Aberrant splicing
12	AT188LA[b]	M	**1348delG**	Termination
12	AT195LA[a]	M	**1348delG**	Termination
15	AT194LA[a]		**2066C>G**	Termination
IVS20	AT195LA[b]	J	IVS20-579delAAGT	Pseudo-exon inclusion
IVS20	AT10LA[a]	J	IVS20-579delAAGT	Pseudo-exon inclusion
IVS20	AT189LA		**IVS20-3del3ins7**	Aberrant splicing
27	AT36LA[b]		3663G>A	Termination
27	AT185LA[h]	T	**3673C>T**	Termination
28	AT196LA[b]	1	3802delG	Termination
IVS31	AT124LA[h]	D	**IVS31-1G>C**	Aberrant splicing
34	AT191LA[h]	S	4777G>T	Aberrant splicing
39	AT194LA[b]		**5566G>T**	Termination
39	AT34LA[a]	3	**5644C>T**	Termination
41	AT187LA[h]	4	5908C>T	Termination
42	AT10LA[b]	K	5932G>T	Termination
IVS53	AT193LA[b]	N	IVS53-2A>C	Aberrant splicing
54	AT155LA[b]		7704_7705delAG	Termination
57	AT192LA[a]	R	**8045C>G**	Thr>Ser (T2682S)
57	AT42LA[b]	I	**8105T>G**	Ile>Arg (I2702R)
57	AT192LA[b]	Q	8150A>C	Aberrant splicing
58	AT161LA[h]	P	**8185C>T**	Termination
64	AT186LA[a]	C	**8874_8877delTGAC**	Termination
64	AT186LA[b]	6	8977C>T	Termination
IVS64	AT34LA[b]	G	IVS64-1G>C	Aberrant splicing

(B)

FIGURE 4. **Haplotypes and mutations of 18 Hispanic-American families. Note heterogeneity, as compared to Spanish and Brazilian families. A: Recurring haplotypes are color shaded, and were assigned numbers instead of letters whenever they were also observed in non-Hispanic-American families. The < > indicate that phase has not been defined. B: Mutations corresponding to affected haplotypes shown in (A). Bolded mutations have not been reported previously. Underlined amino acids are conserved. Superscripts: a, first allele; b, second allele; h, homozygote. Nucleotide numbering is based on +1 being the A of the first translational start codon.**

TABLE 1. **Standardized Costa Rican Haplotypes, with previously published mutations [Telatar et al., 1998a]**

Haplotype	Mutation	S1819	NS22	S2179	S1818
[4]	5908C>T	131	173	143	160
[B]	IVS63del17kb	131	171	141	160
[C]	7449G>A	137	163	139	160
[D]	4507C>T	139	161	141	160
[E]	8264_8268delATAAG	131	169	141	162
[F]	1120C>T	131	169	141	146

occurred on different STR and SNP haplotypes, they most likely represent independent mutational events or true hot spots (see the Discussion section).

Masked Mutations

Three masked mutations were identified that would be missed by conventional PCR-based screening methods: 1) IVS28+1711del3450 (SPAT 5.3, BRAT 10, 13, 16.3, 24, AT162LA); 2) IVS63del~17kb (BRAT 3); and 3) IVS20-579delAAGT (AT195LA, AT10LA). The IVS28+1711del3450 mutation, carried on haplotype IBERIAN [2], was missed when heterozygous genomic DNA was screened by SSCP and DHPLC; it was only identified in retrospect after it was detected in a homozygous Brazilian patient (BRAT 24), carrying the same haplotype (IBERIAN [2]), on whom no PCR product could be obtained for the deleted exons (for further details see Fig. 3). Similarly, IVS63del~17kb, found on Brazilian and Costa Rican patients, would be missed in heterozygous patients. IVS20-579delAAGT, in contrast, was missed on screening because PCR primers did not include the deep intronic region of the mutation. The inclusion of a pseudo-exon was detected in the cDNA.

DISCUSSION

Of the 62 different Iberian A-T mutations identified, 28 have not been reported previously. Five of 110 alleles could not be identified by PTT, SSCP, and DHPLC. Since these alleles correspond to heterozygous patients, they may consist of large genomic deletions that would not be detected by PCR-based strategies unless observed in homozygotes. The high frequency of recurring haplotypes in the Spanish and Brazilian families (65%), and in the previously published Costa Rican families (>95%), stands in sharp contrast to the eight of 18 Hispanic-American families with recurring haplotypes (44%). The marked diversity of affected A-T haplotypes in the Hispanic-American population was 78% (28 of 36 haplotypes); this diversity has been observed in other studies as well.

In a previous study, splicing mutations accounted for approximately half of the mutations in A-T patients [Teraoka et al., 1999], as compared to 26% in our data. They typically involve the highly conserved splice donor (5′ or GT) or acceptor (3′ or AG) sites. We detected one mutation in BRAT 14 (IVS54–3T>G) that affected the

TABLE 2. ATM Mutations on STR and SNP Haplotypes in Patients of Iberian Origin

Mutation	STR haplotype	Brazilian	Spanish	Hispanic-American	Costa Rican	SNP haplotype
A. Identical mutations on *identical* STR and SNP haplotypes (number of families)						
3802delG	IBERIAN [1]	(3)		(1)		H2
IVS28+1711del3450	IBERIAN [2]	(5)	(1)			H2
5644C>T	IBERIAN [3]	(1)	(1)	(1)		H2
5908C>T	IBERIAN [4]			(1)	(3)	H4
B. Identical mutations on *different* STR and/or SNP haplotypes (number of families)						
IVS21+1G>A	SPAT [E]		(1)			H2
	SPAT [G]		(1)			H2
8264_8268delATAAG[a]	IBERIAN [5]	(3)	(1)			H2
	CRAT [E]				(1)	H4
8977C>T[a]	IBERIAN [6]		(3)	(1)		H2
	SPAT [K]		(2)			H3
IVS63del~17kb	BRAT [C]	(1)				H4
	CRAT [B]				(2)	H4

[a]Mutation detected on different STR and SNP haplotypes.

less well conserved position –3 in the acceptor site of intron 54, causing the deletion of exon 55. Another splicing mutation, IVS20-579delAAGT, is identical to one described as a deletion of 4 nt (GTAA) in intron 20 by Pagani et al. [2002]. This mutation abolishes the interaction of the U1 snRNP with DNA causing the insertion of a pseudo-exon in the mRNA, and defining a new splicing motif: the "intron-splicing processing element (ISPE)."

Missense mutations in the *ATM* gene have been associated with an increased risk of breast and other cancers [Gatti et al., 1999; Spring et al., 2002; Concannon et al., 2002]. The three missense mutations we detected in the Iberian population (one in Spain and two in the Hispanic-American group) were found within the kinase domain and affected highly conserved amino acids (mouse and pufferfish). These are being tested for function in a mutagenesis assay [Scott et al., 2002].

In the Spanish cohort, one of the novel mutations that was identified (9170_9171delGA) alters the final stop codon of the gene and replaces it with phenylalanine plus three additional amino acids. The SPAT [A] haplotype in Spanish families was always found in a homozygous state, implying consanguinity in each of three families; this haplotype appears to be associated with Spanish gypsies. In the Brazilian population, we observed mutations of various European origins: Spanish (5644C>T, 8264_8268delATAAG), German (3802delG), and Italian (7517_7520delGAGA). These findings corroborate other reports on the ancestry of the Brazilian population [Alves-Silva et al., 2000; Carvalho-Silva et al., 2001]. The African contribution was also corroborated in the Brazilian A-T population by the presence of an African polymorphism in BRAT 3 (5793T>C), which was described previously by Thorstenson et al. [2001]. Within the Hispanic-American population, two unrelated patients carried the 103C>T founder mutation of North African Jews, and shared STR and SNP haplotypes. Haplovariants associated with this mutation in other Sephardic Jews were reported by Campbell et al. [2003].

In general, identical mutations were found on identical STR and SNP haplotypes; however, we found four exceptions to this rule: IVS21+1G>A, 8264_8268delATAAG, 8977C>T, and IVS63del~17kb (Table 2B). Even the SNP haplotypes were different for 8264_8268delATAAG and 8977C>T, suggesting that these represent independent mutational events or true hot spots in the *ATM* gene. A mechanism for such genomic instability would not be difficult to envision for the IVS63del~17kb, which deletes genomic material between two very homologous LINE-1 repeats. The IVS21+1G>A mutation associated with different STR haplotypes probably reflects the instability of STR markers. In contrast, the mechanisms for why 8264_8268delATAAG and 8977T>C might be hot spots are unclear at this writing.

Through the very efficient use of global haplotype identification with absolute sizes for the STR markers, we have demonstrated that many haplotypes are repeatedly observed within the Iberian population, and that these can help to identify specific ATM mutations. By haplotype prescreening, we have achieved the highest mutation detection rates of any published work to date ($\geq$94%). For this, it is important to emphasize that global STR haplotyping can only be accomplished if all alleles are standardized in advance of testing.

Standardized haplotyping has also allowed several masked mutations to be identified in patients once the mutation was linked to those haplotypes in homozygous patients. This ever-expanding group of masked ATM mutations may be of importance when searching for ATM mutations associated with various malignancies. In order to detect these mutations, specific PCR primers for each would have to be designed and included in such screening. This is especially important when screening DNA for large cancer cohorts for heterozygous mutations. Prescreening with standardized STR markers may also be useful for restudying A-T patients described in previous publications for whom mutations could not be identified.

ACKNOWLEDGMENTS

The authors thank Drs. Ellen Dantos, Teresa Español, Juan Llerena, Montserrat Mila, and Salmo Raskin for providing blood samples on some of the patients described.

REFERENCES

Alves-Silva J, Santos M, Guimaraes PEM, Ferreira ACS, Bandelt H-J, Pena SDJ, Prado VF. 2000. The ancestry of Brazilian mtDNA lineages. Am J Hum Genet 67:444–461.

Athma P, Rappaport R, Swift M. 1996. Molecular genotyping shows that ataxia-telangiectasia heterozygotes are predisposed to breast cancer. Cancer Genet Cytogenet 92:130–134.

Bonnen PE, Story MD, Ashorn CL, Buchholz TA, Weil MM, Nelson DL. 2000. Haplotypes at ATM identify coding-sequence variation and indicate a region of extensive linkage disequilibrium. Am J Hum Genet 67:1437–1451.

Campbell C, Mitui M, Eng L, Coutinho G, Thorstenson Y, Gatti R. 2003. ATM mutations on distinct SNP/STR haplotypes in ataxia-telangiectasia patients of differing ethnicities reveals ancestral founder effects. Hum Mutat 21:123–131.

Carvalho-Silva DR, Santos FR, Rocha J, Pena SDJ. 2001. The phylogeography of Brazilian Y-chromosome lineages. Am J Hum Genet 68:281–286.

Castellvi-Bel S, Sheikhavandi S, Telatar M, Tai LQ, Hwang M, Wang Z, Yang Z, Cheng R, Gatti RA. 1999. New mutations, polymorphisms, and rare variants in the ATM gene detected by a novel SSCP strategy. Hum Mutat 14:156–162.

Chenevix-Trench G, Spurdle AB, Gatei M, Kelly H, Marsh A, Chen X, Donn K, Cummings M, Nyholt D, Jenkins MA, Scott C, Pupo GM, Dork T, Bendix R, Kirk J, Tucker K, McCredie MR, Hopper JL, Sambrook J, Mann GJ, Khanna KK. 2002. Dominant negative ATM mutations in breast cancer families. J Natl Cancer Inst 94:205–215.

Concannon P, Gatti RA. 1997. Diversity of ATM gene mutations detected in patients with ataxia-telangiectasia. Hum Mutat 10:100–107.

Concannon P. 2002. ATM heterozygosity and cancer risk. Nat Genet 32:89–90.

Dork T, Bendix R, Bremer M, Rades D, Klopper K, Nicke M, Skawran B, Hector A, Yamini P, Steinmann D, Weise S, Stuhrmann M, Karstens JH. 2001. Spectrum of ATM gene mutations in a hospital-based series of unselected breast cancer patients. Cancer Res 61:7608–7615.

Easton DF. 1994. Cancer risk in A-T heterozygotes. Int J Radiat Biol 66:S177–S182.

Gatti RA, Tward A, Concannon P. 1999. Cancer risk in ATM heterozygotes: a model of phenotypic and mechanistic differences between missense and truncating mutations. Mol Genet Metab 68:419–423.

Gatti RA, Becker-Catania S, Chun H, Sun X, Mitui M, Lai C-H, Khanlou N, Babaei M, Cheng R, Clark C, Huo Y, Udar N, Iyer R. 2001. The pathogenesis of ataxia-telangiectasia. Clin Rev Allergy Immunol 20:87–108.

Gilad S, Bar-Shira A, Harnik R, Shkedy D, Ziv Y, Khosravi R, Brown K, Vanagaite L, Xu G, Frydman M, Lavin MF, Hill D, Tagle DA, Shiloh Y. 1996. Ataxia-telangiectasia: founder effect among North African Jews. Hum Mol Genet 5: 2033–2037.

Laake K, Telatar M, Geitvik GA, Hansen RO, Heiberg A, Andresen AM, Gatti RA, Borresen-Dale A-L. 1998. Identical mutation in 55% of the ATM alleles in 11 Norwegian AT families: evidence for a founder effect. Eur J Hum Genet 6: 235–244.

Pagani F, Buratti E, Stuani C, Bendix R, Dork T, Baralle FE. 2002. A new type of mutation causes a splicing defect in ATM. Nat Genet 30:426–429.

Rotman G, Vanagaite L, Collins FS, Shiloh Y. 1994. Three dinucleotide repeat polymorphisms at the ataxia-telangiectasia locus. Hum Mol Genet 3:2079.

Savitsky K, Sfez S, Tagle DA, Ziv Y, Sartiel A, Collins FS, Shiloh Y, Rotman G. 1995. The complete sequence of the coding region of the ATM gene reveals similarity to cell cycle regulators in different species. Hum Mol Genet 4:2025–2032.

Scott SP, Bendix R, Chen P, Clark R, Dork T, Lavin MF. 2002. Missense mutations but not allelic variants alter the function of ATM by dominant interference in patients with breast cancer. Proc Natl Acad Sci 99:925–930.

Shiloh Y, Kastan MB. 2001. ATM: genome stability, neuronal development, and cancer cross paths. Adv Cancer Res 83: 209–254.

Sommer SS, Buzin CH, Jung M, Zheng J, Liu Q, Jeong SJ, Moulds J, Nguyen VQ, Feng J, Bennett WP, Dritschilo A. 2002. Elevated frequency of ATM gene missense mutations in breast cancer relative to ethnically matched controls. Cancer Genet Cytogenet 134:25–32.

Spring K, Ahangari F, Scott SP, Waring P, Purdie DM, Chen PC, Hourigan K, Ramsay J, McKinnon PJ, Swift M, Lavin MF. 2002. Mice heterozygous for mutation in Atm, the gene involved in ataxia-telangiectasia, have heightened susceptibility to cancer. Nat Genet 32:185–190.

Stankovic T, Kidd AM, Sutcliffe A, McGuire GM, Robinson P, Weber P, Bedenham T, Bradwell AR, Easton DF, Lennox GG, Haites N, Byrd PJ, Taylor AM. 1998. ATM mutations and phenotypes in ataxia-telangiectasia families in the British Isles: expression of mutant ATM and the risk of leukemia, lymphoma, and breast cancer. Am J Hum Genet 62:334–345.

Swift A, Morrell D, Massey RB, Chase CL. 1991. Incidence of cancer in 161 families affected by ataxia-telangiectasia. N Engl J Med 325:1831–1836.

Taylor AMR, Byrd PJ, McConville CM, Thacker S. 1994. Genetic and cellular features of ataxia-telangiectasia. Int J Radiat Biol 65:65–70.

Telatar M, Wang Z, Udar N, Liang T, Bernatowska-Matuszkiewicz E, Lavin M, Shiloh Y, Concannon P, Good RA, Gatti RA. 1996. Ataxia-telangiectasia: mutations in ATM cDNA detected by protein-truncation screening. Am J Hum Genet 59:40–44.

Telatar M, Wang Z, Castellvi-Bel S, Tai L-Q, Sheikhavandi S, Regueiro JG, Porras O, Gatti RA. 1998a. A model for ATM heterozygote identification in a large population: four founder-effect ATM mutations identify most of Costa Rican patients with ataxia telangiectasia. Mol Genet Metab 64:36–43.

Telatar M, Teraoka S, Wang Z, Chun HH, Liang T, Castellvi-Bel S, Udar N, Borresen-Dale AL, Chessa L, Bernatowska-Matuszkiewicz E, Porras O, Watanabe M, Junker A, Concannon P, Gatti RA. 1998b. Ataxia-telangiectasia: identification and detection of founder-effect mutations in the ATM gene in ethnic populations. Am J Hum Genet 62: 86–97.

Teraoka S, Telatar M, Becker-Catania S, Liang T, Onengut S, Tolun A, Chessa L, Sanal O, Bernatowska E, Gatti RA, Concannon P. 1999. Splicing defects in the ataxia-telangiectasia gene, ATM: underlying mutations and consequences. Am J Hum Genet 64:1617–1631.

Thorstenson YR, Shen P, Tusher VG, Wayne TL, Davis RW, Chu G, Oefner PJ. 2001. Global analysis of ATM polymorphism reveals significant functional constraint. Am J Hum Genet 69:396–412.

Udar N, Farzad S, Tai LQ, Bay J-O, Gatti RA. 1999. A highly polymorphic complex microsatellite marker within the ATM gene. Am J Hum Genet 82:287–289.

Uhrhammer N, Lange E, Porras O, Naeim A, Chen X, Sheikhavandi S, Chiplunkar S, Yang L, Dandekar S, Liang T, Patel N, Teraoka S, Udar N, Calvo N, Concannon P, Lange K, Gatti RA. 1995. Sublocalization of an ataxia-telangiectasia gene distal to D11S384 by ancestral haplotyping in Costa Rican families. Am J Hum Genet 57:103–111.

Vanagaite L, James MR, Rotman G, Savitsky K, Bar-Shira A, Gilad S, Ziv Y, Uchenik V, Sartiel A, Collins FS, Sheffield VC, Richard III CW, Weissenbach J, Shiloh Y. 1995. A high-density microsatellite map of the ataxia-telangiectasia locus. Hum Genet 95:451–454.