Am. J. Hum. Genet. 59:320–330, 1996

Mutations Associated with Variant Phenotypes in Ataxia-Telangiectasia

Carmel M. McConville,[1,*] Tatjana Stankovic,[1] Philip J. Byrd,[1] Germaine M. McGuire,[1] Qing-Yun Yao,[1] Graham G. Lennox,[2] and A. Malcolm R. Taylor[1]

[1]CRC Institute for Cancer Studies, The Medical School, University of Birmingham, Birmingham; and [2]Division of Clinical Neurology, Queens Medical Centre, Nottingham

Summary

We have identified 14 families with ataxia-telangiectasia (A-T) in which mutation of the ATM gene is associated with a less severe clinical and cellular phenotype (~10%–15% of A-T families identified in the United Kingdom). In 10 of these families, all the homozygotes have a 137-bp insertion in their cDNA caused by a point mutation in a sequence resembling a splice-donor site. The second A-T allele has a different mutation in each patient. We show that the less severe phenotype in these patients is caused by some degree of normal splicing, which occurs as an alternative product from the insertion-containing allele. The level of the 137-bp PCR product containing the insertion was lowest in two patients who showed a later onset of cerebellar ataxia. A further four families who do not have this insertion have been identified. Mutations detected in two of four of these are missense mutations, normally rare in A-T patients. The demonstration of mutations giving rise to a slightly milder phenotype in A-T raises the interesting question of what range of phenotypes might occur in individuals in whom both mutations are milder. One possibility might be that individuals who are compound heterozygotes for ATM mutations are more common than we realize.

Introduction

The inherited disorder ataxia-telangiectasia (A-T), is characterized typically by a progressive cerebellar ataxia, oculocutaneous telangiectasia, hypersensitivity to ionizing radiation, and predisposition to lymphoid tumors (Sedgwick and Boder 1991). Although the diagnosis is unambiguous in the majority of patients, there is, nevertheless, a considerable degree of variation in both the severity of these features and the expression of a number of additional A-T–associated phenotypes. Severe sinopulmonary infections, for example, are a major problem for some patients, but not for others (Sedgwick and Boder 1991). Similarly, despite the fact that the overall incidence of cancer in A-T patients is only ~10%–15%, the occurrence of a tumor in more than one sibling in a significant number of families and concordance of tumor type in these cases (Spector et al. 1982; Taylor et al. 1996*b*) suggests that cancer predisposition may be greater for some patients than for others.

The most likely explanation for these observations is genetic heterogeneity, either (*a*) as a result of the existence of additional "A-T" genes or (*b*) within a single A-T gene due, for example, to an effect of the position of the mutation within the gene or to the type of mutation involved (e.g., missense mutation vs. large deletion). Although the existence of additional genes is a possibility that has received considerable interest in view of the existence of four A-T complementation groups, linkage data have not, in general, provided any significant support for this possibility (McConville et al. 1994). Furthermore, the recent identification of a single ATM gene that is mutated in patients from all four complementation groups (Savitsky et al. 1995*a*) suggests that this gene is responsible for the A-T phenotype in the majority of patients. Many different mutations have recently been identified in A-T patients (Savitsky et al. 1995*a*; Byrd et al. 1996; Gilad et al. 1996), the majority of these being null mutations (Gilad et al. 1996).

The cloning of the ATM gene and the availability of sequence data for the complete coding region (Savitsky et al. 1995*b*; Byrd et al. 1996) now provides the opportunity to determine the relationship between the clinical/cellular phenotype and the nature of the associated mutation(s). This information may also provide some clues as to the location of functionally important domains within this relatively large gene (~9 kb of coding sequence). At present, the precise function(s) of the ATM

Received February 23, 1996; accepted for publication May 22, 1996.

Address for correspondence and reprints: Dr. Malcolm R. Taylor, CRC Institute for Cancer Studies, The University of Birmingham, Birmingham B15 2TJ, United Kingdom. E-mail: tayloramr@vax1.bham.ac.uk

*Present address: Institute of Child Health, University of Birmingham, Birmingham.

0002-9297/96/5902-0007$02.00

protein is unknown, with only a portion of the COOH-terminal region showing any homology to other proteins of known function (Savitsky et al. 1995*a*). This homology is of interest, however, since it suggests that the ATM protein belongs to a family of proteins, all of which contain a phosphatidylinositol 3-kinase (PI3-Kinase) motif at their COOH-termini and many of which appear to play a role in one or more cellular phenotypes that are defective in A-T cells. The yeast TOR proteins, for example, as well as their mammalian homologues FRAP and rRAFT, are required for transition between G_1 and S phase (Brown et al. 1994; Sabatini et al. 1994), while MEC1 (*Schizosaccharomyces cerevisiae*) and rad 3 (*S. pombe*) are both required for S→M and G_2→M cell cycle checkpoints. Rad3 is also required for DNA repair (Jimenez et al. 1992; Seaton et al. 1992). It has also been suggested that the Drosophila gene mei-41 is a structural and functional homologue of the ATM gene, in view of sequence similarity to the ATM gene and also the radiation sensitivity and chromosome instability that result from mutations in mei-41 gene (Hari et al. 1995).

It is not clear from these studies whether, or how, mutations in a putative PI3-K domain might influence aspects of the A-T phenotype such as neurological function. It has been shown, however, that members of the PI3-kinase family differ considerably both from each other and from the A-T gene in the NH_2-terminal region, perhaps as a consequence of additional different functions or interactions. Therefore, it is important to determine to what extent mutations in different parts of the gene might influence different aspects of the phenotype.

As part of a larger study to investigate the relationship between genotype and phenotype in A-T, we have examined a distinct A-T subgroup in which the overall phenotype appears to be less severe at both the clinical level and the cellular level, than in the majority of patients. This subgroup includes ~15% of all U.K. A-T patients (Taylor et al. 1994, 1996*a*). In these variant patients, the ataxia shows a later mean age at onset and a slower rate of neurological deterioration. In addition, chromosomal radiosensitivity is either significantly reduced or absent (Taylor et al. 1994, 1996*a*). Our previous studies showed the presence of the same chromosome 11 haplotype, over the region of the ATM gene, in eight variant patients (designated as "haplotype 4-1-8," on the basis of the presence of the alleles 4, 1, and 8, respectively for the three core loci D11S1817, D11S1343, and D11S1819) (Taylor et al. 1994). We concluded that the patients in these families carry an identical mutation on one copy of chromosome 11 and that this is likely to be in the ATM gene. We also speculated that this represents a "mild" mutation, which may have a mitigating effect on a more "severe" mutation on the second copy of chromosome 11. In the present paper, we describe the molecular basis of the mutations responsible for variant phenotypes in patients both with and without the conserved 4-1-8 haplotype and show that at least two different mechanisms exist whereby a milder cellular and clinical phenotype may be produced.

Patients and Methods

Patients

A total of 18 A-T patients from 14 families were studied; their clinical features are shown in table 1. Their parents noted a later mean age at onset than those of patients with "classical" A-T (as judged on both clinical and cell biological grounds), i.e., 3.8 years compared to 1.5 years in 31 consecutive classical patients assessed at the U.K. A-T clinic in Nottingham. Comparing only adult patients (i.e., those aged ≥16 years), the 14 variant patients are less severely affected in all their neurological features than the 10 consecutive classical patients assessed in this clinic, despite being well matched for age. This implies that the variant patients have a slower rate of progression. None of the variant patients has had problems with serious infection, leukemia, or lymphoma.

The patients included 12 (from 10 families) with the 4-1-8 haplotype (table 1A) and 6 others (in 4 families) without this haplotype (table 1B). All these patients were initially classified as having a variant A-T phenotype on the basis of reduced radiosensitivity (i.e., chromosomal radiosensitivity was low, and colony-forming ability of fibroblasts from all patients tested, following exposure of the cells to gamma-rays, was intermediate between normal and classical A-T patients). This intermediate radiosensitivity has been reported previously for 14 of the 18 patients in the present study (Taylor et al. 1996*a*).

Haplotype Analysis

Genotyping of A-T individuals and their families was carried out for the loci listed in table 2. Primer sequences used for the amplification of short tandem repeat (STR) polymorphisms were kindly provided by Y. Shiloh (S1816, S1817, S1819, S2179, S2180, S1818, S2178, and S1960, also listed in the Genome Data Base [GDB]). Primer sequences used for the remaining loci were as listed in GDB, with the exception of S2000 and S2004, for which modified primers were used: S2000: (*a*) 5′-CAA GCA TGA CTT CAT ATA TAG ACC and (*b*) 5′-GAA GAT CTG TGA AAT GTG CCA G; S2004: (*a*) 5′- GCT TCT CAG CCT GCA TAT CTG and (*b*) 5′-AGG CTG TTT GCT GGG TTA TGG. One primer for each locus was synthesized with a fluorescent label (FAM, HEX, or TET), PCR products were electrophoresed on a ABI Genescan System, and data were processed using ABI Genotyper Fragment Analysis software. Allele number assignments are as listed in GDB. (It should be noted that the locus S2004 is incorrectly listed as S2003 in some publications.)

Table 1

Clinical Features of A-T Patients with a Variant Cellular and Clinical Phenotype

Patient	Age (years)	Sex	Onset (years)	Gait Ataxia	Limb Ataxia	Dysarthria	Chorea and/or Dystonia	Eye Movements	Peripheral Neuropathy	Infection
A. Patients sharing the 418 mutation:										
1-3	26	F	8	2	2	1	1	2	0	0
14-4	18	F	3	2	1	1	1	1	1	1
38-3	30	M	3	2	1	1	1	1	2	0
38-4	28	M	3	2	1	1	1	1	2	0
40-3	28	M	5	3	1	2	0	2	2	0
40-4	24	M	2	2	1	2	1	2	2	2
44-4	16	F	2.5	1	1	1	1	1	0	0
45-3	26	M	2.5	3	2	1	0	1	2	0
52-4	35	F	1	2	2	1	1	3	2	0
59-4	12	F	8	1	1	1	1	1	1	0
62-4	19	F	12	1	1	1	2	1	3	0
79-3	39	M	1.5	4[a]	2	1	1	1	1	0
B. Patients with different mutations:										
46-3	25	M	4	2	1	2	0	2	1	2
46-4	14	M	8	1	1	1	1	1	0	2
53-3	21	M	1[b]	2	2	2	1	2	0	0
53-4	18	M	1[b]	2	1	2	2	3	1	0
78-3	7	F	1.5	1	0	1	1	0	0	0
88-3	2.5	F	1.5	1	0	0	0	0	0	0

NOTE.—Gait ataxia: 0 = none; 1 = mild, walks unaided; 2 = moderate, stands unaided, walks with frame or assistance of one; 3 = severe, stands and walks <10 m with assistance of one or two; 4 = very severe, cannot stand, wheelchair dependent. Limb ataxia, dysarthria, chorea and/or dystonia: 0 = none; 1 = mild; 2 = moderate; 3 = severe. Eye movements: 0 = normal; 1 = mild abnormalities: gaze evoked nystagmus, broken pursuit, hypometric saccades; 2 = moderate abnormalities, including increased latency and reduced velocity of saccades; 3 = severe abnormalities, including oculomotor apraxia; 4 = very severe abnormalities, including ophthalmoplegia. Peripheral neuropathy: 0 = none; 1 = mild, with absent ankle jerks and impaired vibration sense in the feet; 2 = moderate, with distal wasting with or without weakness, widespread areflexia and impaired pain sensation in the feet; 3 = severe, with disabling wasting and weakness. Infections: 0 = no excess; 1 = possible excess; 2 = definite excess or one severe infection; 3 = recurrent severe infections.

[a] Grade 3 (walking with assistance) when aged 38 years.

[b] Presented with chorea and gait ataxia. Onset was with gait ataxia in all other cases.

Mutation Detection

Mutation detection and sequencing were carried out as described elsewhere (Byrd et al. 1996). The primers used for genomic and cDNA sequencing of the 4-1-8 mutation (fig. 1) are as follows: 8498: 5′-TCA GAG CAC TTT TTC CGA TGC TG; 8421: 5′-CTT CTC CAA CTC CTG CTA ATG TC; 8451: 5′-CCA TTT TAA TGC TGA ATA AGA TCC TG; and 8152: 5′-AAG TGA GCA GCA CAA GAC TGA GC.

The pair of primers used for allele-specific PCR giving amplification from the 4-1-8 allele only in patient 62-4, who also carried a 7636del9 mutation, were 5′-primer 8498, as above, and 9977: 5′-ATG GGG GTG ATC CAT TGA AAT TCT AG.

Hybridization and Phosphorimager Quantitation Analysis

The blot containing exon-exon reverse-transcriptase (RT)–PCR products from 4-1-8 and non–4-1-8 alleles was hybridized with a total PCR product amplified with primers 7873 and 7874 from the two exons 36 and 43, respectively (7873: 5′-CTT CAG TGG ACC TTC ATA ATG C and 7874: 5′-CCA TAC AAA CTA TCT GGC TCC). The blot containing exon/insert RT-PCR products was hybridized with a total exon/insert RT-PCR product amplified with primers 8498 and 8421. The image was scanned on a Molecular Dynamics (SI) phosphorimager.

Protein-Truncation Test (PTT)

A modified primer, containing the T7 promoter and a eukaryotic translation–initiation sequence, was used to generate PCR products suitable for PTT analysis (primer 8650: 5′-GGA TCC TAA TAC GAC TCA CTA TAG GAA CAG ACC ACC ATG AAC ACA CTG GTA GAA GAT TGT GTC, in combination with primer 8651: 5′-AGT ATC TGA AAA CCG GGC TAA TGA

Table 2

Chromosome 11q22-q23 Extended Haplotypes in Variant AT Patients

	Family No.									
Locus	1	14	38	40	44	45	52	59	62	69
D11S1816 7L	7	7	7	7	7	7	7	7	6	...
D11S2000 GATA28D01	...	11	11	11	8	11	8	11	9	...
D11S1817 A1	4	4	4	4	4	4	4	4	4	4
D11S1343 AFM296YD9	1	1	1	1	1	1	1	1	1	1
D11S1787 AFM152WH9	3	3	1	3	3	3	3	...	3	3
D11S1819 A4	8	8	8	8	8	8	8	8	8	8
	A-T									
D11S2179 H2	4	4	4	4	4	5	4	...	4	5
D11S1778 AFM047XC5	2	1	2	2	1	2	2	...	12	2
D11S1294 UT928	1	4	1	4	4	1	1	2	2	...
D11S2180 67-2	4	4	4	4	4	2	4	4	4	4
D11S1818 A2	1	1	1	1	1	3	1	1	1	1
D11S2178 67-4	5	7	7	7	6	4	5	7	...	7
D11S2004 GATA3C07	3	8	1	2	11	3	9	...	2	3
D11S1960 409-1	1	2	4	1	1	2	1	4	3	...
D11S927 AFM225YB4	7	5	5	3	11	5	11	1	11	1

Note.—Loci are listed in chromosome order, from centromere to telomere. The position of the ATM gene relative to the markers is shown. The locus D11S2179 is located at the extreme 3′ (telomeric) end of the ATM gene. The distance between markers D11S1816 and D11S927 is ~5 Mb. The majority of these patients had in common alleles 4, 1, and 8 for markers D11S1817, D11S1787, and D11S1819, respectively. Underlined figures indicate alleles common to different patients across the haplotype. The haplotypes show a high degree of similarity over a distance of ~3.5 Mb between the loci D11S1816 and D11S2178. Patients listed were in a group initially recognized as showing an intermediate level of cellular and chromosomal radiosensitivity, a slower rate of progress of the disorder, and the absence of both serious infection and either lymphoma or leukemia.

GAG). This fragment corresponds to nucleotide position 5157-7188 of the A-T cDNA. PTT analysis was carried out using a TNT Coupled Reticulocyte Lysate System as recommended by the manufacturer (Promega). ^{35}S-labeled protein products were separated on a 13% SDS-polyacrylamide gel and detected by autoradiography.

Results

Haplotypes in the Region of the A-T Gene

Previous studies, which identified a common (4-1-8) haplotype on one of the two A-T chromosomes in eight low-radiosensitivity families (Taylor et al. 1994), were extended with the inclusion of two additional families and the analysis of a further eight STR loci. The resulting haplotypes (table 2) show a high degree of similarity over a distance of ~3.5 Mb between the loci D11S1816 and D11S2178, thus confirming a probable founder effect in this group of families. This interval includes the locus D11S2179 which is located at the extreme 3′ end of the ATM gene (Savitsky et al. 1995*a*). The remaining four variant families did not show any evidence of haplotype similarity to each other or to the 4-1-8 group, suggesting that different mutations are involved.

Identification of A-T Gene Mutations

The complete coding sequence of the ATM gene (9,168 bp; Savitsky et al. 1995*b*; Byrd et al. 1996) was amplified in a series of eight fragments (designated 5′ VI, VII, VIII, II, I, III, V, IV 3′) using an RT-PCR strategy. Mutation identification was undertaken by use of the techniques of heteroduplex analysis and restriction endonuclease fingerprinting (Liu and Sommer 1995), followed by sequencing of the appropriate region of the gene. A total of 10 different mutations was identified, and these were present on 19/28 (68%) A-T chromosomes (no mutations were detected on 9 chromosomes) (table 3). With one exception, each mutation was represented only once, and all patients studied were compound heterozygotes.

A Splicing Mutation Common to 10 A-T Families

As predicted from previous studies, all patients with the 4-1-8 haplotype in the region extending from D11S1816 to D11S2178 on one A-T chromosome had

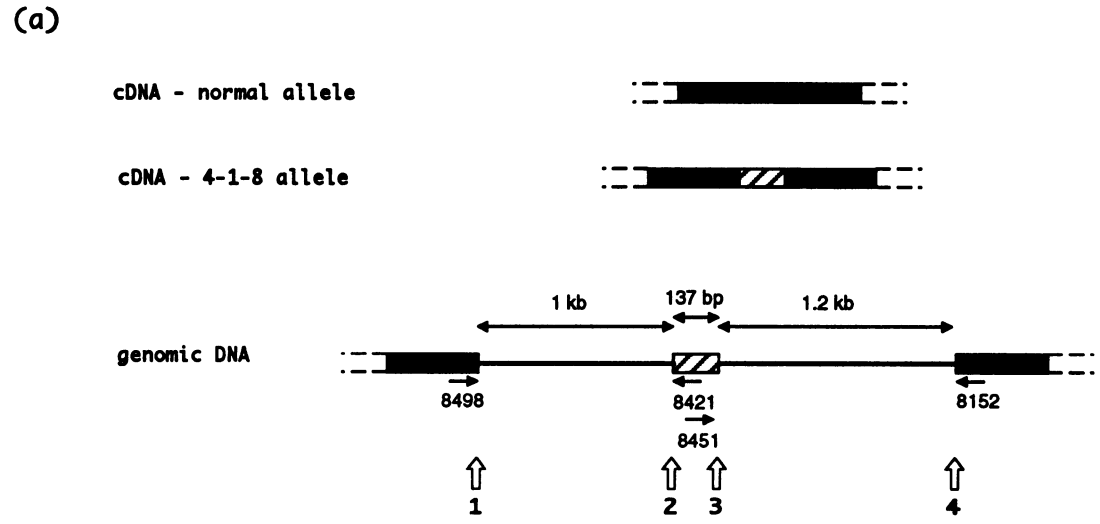

Figure 1 *a*, cDNA and genomic DNA structure in the region of the 4-1-8 insertion. Exons are indicated by solid boxes and introns by horizontal lines. The insertion is indicated by a hatched box. The positions of PCR primers 8498, 8421, 8451, and 8152 are shown. Open vertical arrows indicate the positions of exon/intron and intron/insertion boundary sequences. *b*, DNA sequence at exon/intron and intron/insertion boundaries. Exon and insertion sequences are shown in uppercase and intron sequence in lowercase letters. Consensus splice-site sequences are shown below each sequence, for comparison. Invariant positions in the splice consensus sequences are shown in boldface type. The A→G mutation in sequence 3 is boxed. *c*, Sequence of the 137-bp insertion. Uppercase letters denote insertion sequence, and lowercase letters denote intron sequence.

an identical mutation. This was detected initially as an additional RT-PCR product in fragment I (position 5001-6059 in the ATM DNA sequence), ~100 bp larger than predicted from the sequence data. This larger fragment was cloned and partially sequenced, revealing an insertion of 137 bp between bases 5762 and 5763. This did not appear to result from the inclusion of an additional coding exon, since multiple stop codons were present in the insertion in all three reading frames. The predicted consequence of this mutation is the production

Table 3

Mutations in Variant A-T Patients

	Mutation 1			Mutation 2		
Patient	Nucleotide	Codon	Protein	Nucleotide	Codon	Protein
A. Patients sharing the 4-1-8 mutation:						
1-3	5762ins137 (exons 40/41)	1921	stop 1930	3801 del G (exon 28)	1268	stop 1268
14-4	5762ins137	1921	stop 1930	n.d.	...	...
38-3, 38-4	5762ins137	1921	stop 1930	n.d.	...	...
40-3, 40-4	5762ins137	1921	stop 1930	8787 ins 14 (exon 62/63)	2930	stop 2933
44-4	5762ins137	1921	stop 1930	n.d.	...	...
45-3	5762ins137	1921	stop 1930	n.d.	...	...
52-4	5762ins137	1921	stop 1930	C 9139 T (exon 65)	3047	stop 3047
59-4	5762ins137	1921	stop 1930	6412 del AG (exon 46)	2139	stop 2144
62-4	5762ins137	1921	stop 1930	7636 del 9 (exon 54)	2546	del SRI
79-3	5762ins137	1921	stop 1930	2282 del CT (exon 17)	762	stop 763
B. Patients with different mutations:						
46-3, 46-4	T7271G (exon 51)	2424	V2424G	n.d.	...	...
53-3, 53-4	n.d.	...	...	n.d.	...	...
78-3	n.d.	...	...	n.d.	...	...
88-3	T8480G (exon 60)	2827	F2827C	1561 del AG (exon 12)	521	stop 521

Note.—n.d. = no mutation detected in A-T gene coding sequence; exon numbering is with respect to exons identified by Uziel et al. (1996).

of a truncated protein including residues 1–1921, and an additional 8 residues from the inserted fragment, thus giving a protein only ~60% of the normal size, and lacking the 1135 C-terminal amino acids (including the putative PI3-kinase domain).

In order to investigate the insertion further, genomic DNA, including this part of the ATM gene, was amplified using primers flanking the insertion point and sequenced to define the genomic structure in normal and 4-1-8 individuals and to look for any mutation in intronic sequence. cDNA-specific primers (8498 and 8152) produced a single genomic product of ~2.2 kb in both normal and 4-1-8 individuals (fig. 1). A comparison of genomic and cDNA sequences indicated that an exon-intron junction is located immediately following position 5762 in the cDNA sequence (i.e., the position of the insertion in the 4-1-8 allele). Intron sequence was identical on 4-1-8 and normal alleles in this region. The same exon-intron boundary sequence was also obtained in normal and 4-1-8 individuals, from a 1-kb PCR product amplified using the 5′ cDNA primer (8498) and a primer corresponding to the 5′ end of the inserted sequence (8421) (fig. 1). The conclusion drawn from these results is that the inserted sequence in the 4-1-8 allele is contained within a 2.2-kb intron and is aberrantly spliced into the A-T transcript in 4-1-8 individuals.

We next sequenced the 3′ intron-exon junction and the regions immediately 5′ and 3′ to the insertion sequence to look for either splice site mutations or evidence of cryptic splice sites that might account for the observed apparently aberrant splicing. No mutation was found in the sequence including the intron-exon boundary. Two significant features were noted, however, in the region flanking the insertion sequence. First, sequences resembling splice-acceptor and splice-donor sites were located in the regions flanking the 5′ and 3′ ends, respectively, of the insertion sequence. Second, a point mutation in the 4-1-8 allele at the +5 position in the splice-donor site improves the match between this sequence and the consensus splice-donor sequence (fig. 1). Since this consensus sequence reflects base pairing with U1 snRNA, any deviation from the consensus, particularly at the most highly conserved positions (+1, +2, and +5; Shapiro and Senapathy 1987), might be expected to affect splicing efficiency.

By use of the scoring system of Shapiro and Senapathy (1987), the splice-donor sites may be ranked according to their match to the consensus sequence: 1: 4-1-8 cryptic site (score: 87.6); 2: true splice site (i.e., flanking the 5′ exon) (81.0); and 3: normal cryptic site (73.2). This provides a possible explanation for the generation of the 4-1-8 transcript: the occurrence of the 4-1-8 mutation favors the splicing of the 3′ exon to the 3′ end of the insertion sequence, followed by splicing of the 5′ exon to the 5′ end of this sequence. In the absence of the 4-1-8 mutation, the 3′ exon is preferentially spliced to the 5′ exon, with the loss of all intervening sequence. Bases 5762 and 5763 of the complete ATM cDNA (Byrd et al. 1996) mark the junction of exons 40 and 41 described by Uziel et al. (1996), thus confirming our obser-

vation that the 137-bp insertion occurs between two exons.

It is surprising that such a major disruption of the ATM protein should be associated with a mild phenotype. It was noted, however, that the RT-PCR product from the 4-1-8 allele was present at a consistently lower level, compared with the product from the second allele in all patients (fig. 2). This could be accounted for by unequal expression from the two alleles. However, an alternative explanation more consistent with the presence of a mild phenotype is that the mutation in the intronic sequence affects the efficiency of correct exon splicing, resulting in the production of both normal and aberrant transcripts from the mutant allele. Phosphorimaging of the blotted RT-PCR showed that, while most patients showed a similar ratio of normal to mutant product, patients 62-4 and 44-4 showed a significantly lower proportion of mutant product, suggesting the presence of some variability in the ratio between

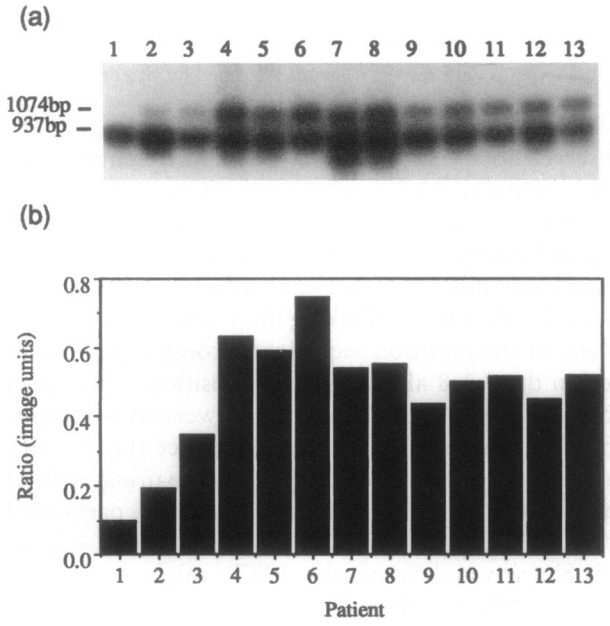

Figure 2 *a,* RT-PCR products obtained by use of primer set I (7873 from exon 36 and 7874 from exon 43). Lane 1 contains a normal control and lanes 2–13 are 4-1-8 patients (patients 62-4, 44-4, 59-4, 1-3, 40-4, 40-3, 14-4, 45-3, 38-3, 38-4, 52-4, and 79-3, respectively). The smaller fragment corresponds to the predicted size from published cDNA sequence (937 bp), and the larger fragment (1074 bp) contains a 137-bp insertion. The variation in band intensity for this larger fragment was seen consistently, following independent PCRs using these and additional primer pairs. *b,* Phosphorimaging profile of the 12 4-1-8 patients, (columns 2–13, patients in the same order as *a*). The profile shows differences between 4-1-8 patients with respect to the ratio of the larger to smaller fragment on image volume with patients 62-4 and 44-4 showing the lowest proportion of 4-1-8 to normal product. The image volume was corrected for background, and the ratio of the image units shows that even in the non-4-1-8 control (column 1) some expression of the larger 4-1-8 product can be detected.

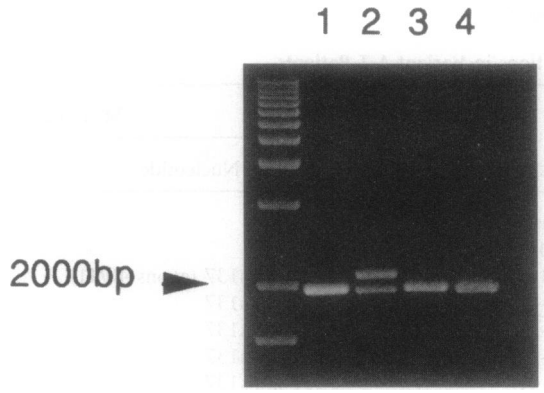

Figure 3 RT-PCR products obtained by use of primers 8498 and 9977. Lane 1 contains a normal control showing only the normal product, and lane 2 is patient 62-4 with both the 4-1-8 mutation and second mutation 7636del9. Both the normal product and the 137-bp insert product are amplified from the 4-1-8 allele alone. Lanes 3 and 4 are two additional unrelated A-T patients (26-3 and 94-3, respectively), without the 4-1-8 mutation but both carrying the 7636del9. These show only the RT-PCR product from the allele without the 9-bp deletion.

these two products in different 4-1-8 patients (figs. 2*a* and 2*b*). Duplicate amplifications and blotting showed the same lower ratio in patients 62-4 and 44-4 (data not shown).

While these results provide strong circumstantial evidence for the role of a splicing error in the 4-1-8–associated phenotype, demonstration of a normal length RT-PCR product and normal protein from the 4-1-8 allele would provide confirmation of this mechanism. We investigated this in two ways. First, using cells from patient 62-4 with the 4-1-8 insertion and a 7636del9 deletion in the second allele (table 3), we carried out RT-PCR using primer 8498 in exon 40, 5′ to the 137-bp insertion and the reverse primer (9977), which included at its 3′ end eight of the nine bases in the 7636del9 deletion. The 3′ end of this primer cannot anneal to an allele with the 7636del9, and therefore no amplification could occur from the allele containing the 9-bp deletion. The expected size of the normal RT-PCR product with these two primers was 1,986 bp. Nevertheless, we obtained two bands differing in size by 137 bp (fig. 3), which indicates the presence of RT-PCR products from both normal and insert-containing transcripts that must both have been derived from the 4-1-8 mutant allele alone in patient 62-4. It is interesting to note that, when only the products from the 4-1-8 allele could be amplified, in patient 62-4, it appeared that the RT-PCR insert product was present at a higher level than the normal sized product. This is in contrast to the ratio of insert to normal RT-PCR products observed when both alleles were amplified (fig. 2, lane 2). By use of the same primers for RT-PCR on a normal control individual and two

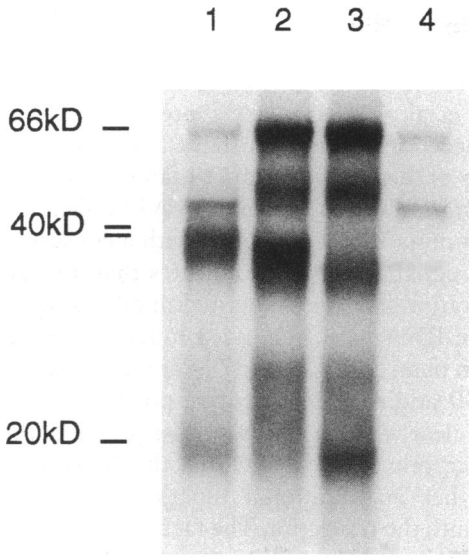

Figure 4 Protein-truncation test in 4-1-8 and non–4-1-8 individuals. Lane 1, 4-1-8 patient 59-4, who is a compound heterozygote for the ATM mutation with both mutations within the amplified fragment analyzed by PTT. A 20-kD protein corresponds to the 4-1-8 allele, and a 40-kD protein to the mutation within the second allele. Protein of the normal size (60–70 kD) can also be detected in this patient, most likely associated with the "leaky" 4-1-8 mutation. Lane 2, Non–4-1-8 patient 5-3 with a different mutation in the same region of the A-T gene as in patient 59-4, giving a protein of just under 40 kd in this PTT. The mutation in the second allele is elsewhere in the A-T gene, and this results in the production of a polypeptide of normal size. Lane 3, 4-1-8 patient 52-4 with only the 4-1-8 mutation in this amplified region, again showing the 20-kD protein from the 4-1-8 allele and a polypeptide of normal size from the second allele. Lane 4, Normal control individual showing the production of normal size protein. The additional band of ~48 kD is rabbit reticulocyte protein present in all lanes.

further A-T patients with the same 7636del9 deletion as patient 62-4, but without the 4-1-8 mutation, only the normal sized product was amplified.

Second, we used the protein-truncation test to look for evidence of a full-length protein product in 4-1-8 patient 59-4 in whom the second ATM mutation was also a truncating mutation again located within the region amplified by the same primers. In addition, we also examined a further 4-1-8 patient (52-4) but with a second mutation outside the region amplified by the primers and a third patient (5-3) with a non–4-1-8 mutation within the region amplified by the primers. Figure 4 shows that, indeed, there is some normal full-length protein produced in cells from patient 59-4 as well as both truncated polypeptides. The normal product is unlikely to have resulted from translation of the non–4-1-8 transcript due to the presence of a frameshift/truncating mutation within coding sequence, in this second allele.

In view of these results, it might be expected that in normal individuals aberrant splicing might occur at a very low level. Although no inserted RT-PCR product was detected in normal individuals by use of exon-specific primers, the use of an exon-specific primer with an insert-derived primer did produce a cDNA product of the predicted size, consistent with this suggestion (fig. 5), and no product of ~1 kb that might be expected from genomic DNA (see fig. 1). In addition, column 1 in figure 2*b* from a normal individual shows evidence for some amplification of a transcript carrying the insertion, consistent with the relatively high Shapiro-Senapathy score for this cryptic site.

The Second Mutation in 4-1-8 Patients

Since it seems likely that the majority of A-T patients are compound heterozygotes, mutations on both A-T chromosomes may influence the overall phenotype. Consequently, we also attempted to identify and characterize mutations in the second A-T chromosome in this group of patients. An additional six mutations were identified in exons 17, 28, 46, 54, 62/63, and 65. Four of the six mutations were small deletions or insertions of 1–14 bp, and these caused frameshifts and were predicted to result in protein truncation (table 3). The second mutation in siblings 40-3 and 40-4 (table 3) was also observed in two additional patients with a classical

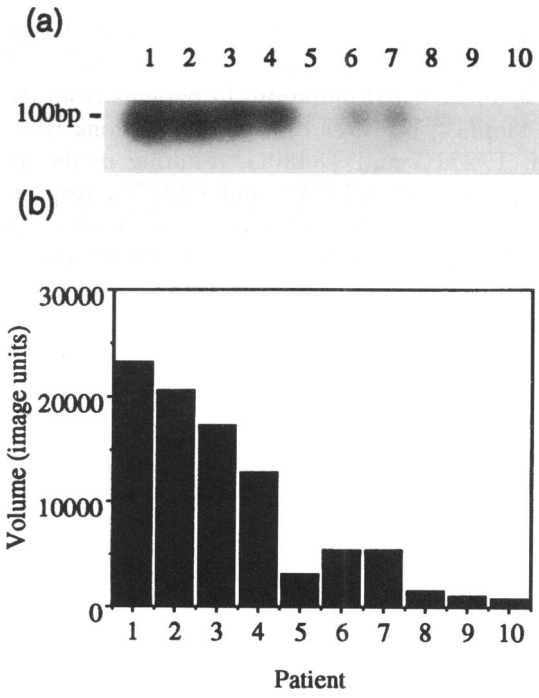

Figure 5 *a*, RT-PCR products of 100 bp, obtained by use of primers 8498 and 8421 from four 4-1-8 individuals (59-4, 1-3, 38-3, and 44-4, lanes 1–4) and six non–4-1-8 individuals (lanes 5–10). *b*, Levels of RT-PCR exon/insert product in both 4-1-8 (columns 1–4, patients as for *a*) and non–4-1-8 individuals (columns 5–10) from phosphorimaging. It is clear that non–4-1-8 individuals also show some expression of this abnormal product.

A-T phenotype (T. Stankovic and C. M. McConville, unpublished results).

One nonsense mutation was identified (C9139T), which produced a stop codon resulting in loss of 10 amino acids from the C-terminal end of the gene. The functional significance of this relatively small deletion is not clear. It is unlikely to represent an irrelevant polymorphism, since the new *Alu*I site that it created was not found on 50 normal chromosomes tested. It may be significant that this deletion includes two amino acids (at positions 3051–3052) which are conserved in many members of the PI3-kinase family (e.g., DNA-PK, FRAP, TOR1, and ESR1).

The only deletion mutation identified that did not result in protein truncation was a 9-bp in-frame deletion involving positions 7636–7644 in exon 54 (amino acids 2546–2548). This is the same mutation reported in three patients of British/Irish origin (Savitsky et al. 1995*a*) and which we have also observed in an additional seven U.K. A-T patients (T. Stankovic and C. M. McConville, unpublished results). The finding of a short in-frame mutation having a severe phenotypic effect may emphasize the functional importance of part or all of the 3′ exon 54.

Missense Mutations Associated with an Altered A-T Phenotype

Of the 14 variant A-T families studied, 4 did not belong to the 4-1-8 haplotype group and did not have the 4-1-8 associated mutation. In patients from two of these families, however, missense mutations were detected, T7271G and T8480G, resulting in the amino acid substitutions V2424G and F2827C, respectively (table 3). This result is of interest, since this type of mutation appears to be uncommon in A-T patients, most of whom have either deletions or insertions that frequently result in a frameshift and protein truncation (Savitsky et al. 1995*a*; Byrd et al. 1996; Gilad et al. 1996; authors' unpublished results). Consequently, the association between a mild phenotype and a mutation that does not cause major disruption of the protein integrity may be significant. Both these mutations are located in the C-terminal 20% of the gene and the latter within the PI3-kinase domain (but not within a conserved amino acid residue). The second mutation has been identified in one of these families and results in a 2-bp deletion leading to truncation and termination of the protein.

Discussion

Our previous studies have shown that a distinct subgroup of A-T patients (comprising ~15% of all U.K. patients) present with a variant phenotype characterized by significantly reduced radiosensitivity, and an apparently milder clinical picture, particularly of the neurological features, as shown for example by a later onset of ataxia and slower rate of neurological deterioration (Taylor et al. 1996*a*). Cancer predisposition may also be reduced or absent. We have now shown that the majority of these patients (10/14 families) share an identical mutation on one of the two ATM alleles. The fact that this mutation is associated with a conserved haplotype (designated "4-1-8") indicates that a single ancestral mutation has been inherited in all cases.

At the RNA level, the 4-1-8 mutation results in the insertion of an additional 137 nt at the junction between exons 40 and 41 (at nucleotide positions 5762/5763). It is not clear whether this is a consequence of the activation of cryptic splice sites within the intervening intron or whether an additional functional exon has been spliced into the transcript. The fact that a range of different ATM transcript sizes (from ~2 kb to 12 kb) has been observed in several different tissues (Savitsky et al. 1995*a*; P. J. Byrd, unpublished results) suggests that alternative exon usage may be important in tissue or developmental stage specific functions of the A-T gene. The amplification of the 4-1-8 insert transcript from normal lymphocyte cDNA and also from a range of tissue specific cDNA libraries (results not shown) is also consistent with this suggestion. Alternative RNA splicing is more usually associated either with the use of different 5′ promoter elements, or with the use of different 3′ polyadenylation signals, rather than with the introduction of a premature stop codon, as was observed in the 4-1-8 transcript. In a recent survey of neurone specific splicing patterns, however, Stamm et al. (1994) listed 11 genes in which alternatively spliced exons encoded stop codons leading to truncated proteins. In most cases, truncation occurred close to the 3′ end of the gene. One exception, however, was noted in the case of the N-methyl-D-aspartate receptor 1 (NMDAR1) gene (Hollman et al. 1993). At least nine splice variants were detected, of which eight involved alternative splicing of exons 5, 6, 21, and 22. The ninth variant was generated through the use of an alternative splice-acceptor site located between exons 2 and 4. This variant, consequently contains an extra exon, exon 3, with an open reading frame of 150 bp, followed by a termination codon. The function of the variant is unknown.

The apparent difference in the relative amounts of transcript produced from each of the two ATM alleles in 4-1-8 patients (fig. 2) may reflect differing levels of transcription from each allele or differences in the stability of the two transcripts. While most 4-1-8 patients showed a similar ratio of normal to mutant cDNA product, there was some variation. In light of the fact that all these patients share the same haplotype and mutation in one allele, i.e., the 4-1-8 mutation, there is no obvious reason why there should be differences between patients

in splicing of RNA transcripts from this allele. Thus, the ratio of the normally spliced transcript and the insertion transcript from the 4-1-8 allele might be expected to be the same in different patients. The observed differences between the levels of the normal and insertion transcripts, in different patients, might be a reflection of the stability of the transcript from the second allele. This might vary according to the nature of the second mutation. The presence of a greater amount of normal sized transcript in some individuals would compete against the larger, insertion transcript in the PCR, giving the appearance of lower expression of the insertion species in those individuals. Although instability of transcripts with premature termination codons has been documented for several genes (Menon and Neufeld 1994; Szabo and King 1995), degradation of the transcript from the 4-1-8 allele would not account for the important fact that this allele acts as a "mild" mutation. A more plausible explanation for the less-severe phenotype is that more than one transcript is produced from the 4-1-8 allele. The occurrence of multiple spliced mRNAs has been reported in other disorders: Lelli et al. (1995), for example, detected three abnormally spliced mRNAs, resulting from a single point mutation at the +1 position of intron 15 in the LDL-receptor gene. It is possible therefore that, in addition to the aberrantly spliced transcript, normal transcript is also produced from the 4-1-8 allele at levels sufficient to moderate the phenotype. More direct evidence in support of this conclusion comes, first, from our demonstration of the presence of two transcripts, one of normal size and one containing the 137-bp insertion, in cells of a patient (62-4) in which only amplification from the 4-1-8 mutant allele could occur and, second, that in vitro translation of mRNA from a 4-1-8 patient (59-4) resulted in the production of normal product together with both predicted truncated polypeptides. The normal product is unlikely to have resulted from translation of the non–4-1-8 transcript, because of the presence of a frameshift/truncating mutation within coding sequence, in this second allele.

The expression of some normal A-T product from the 4-1-8 allele in most or all tissues is consistent with an overall milder clinical and cellular phenotype in 4-1-8 patients (e.g., a slower rate of progress of the cerebellar degeneration, the absence of serious respiratory infections, the presence of a smaller increase in radiosensitivity in both lymphocytes and fibroblasts, and fewer translocations in lymphocytes). Cancer predisposition may also be decreased: since ~15% of A-T patients develop leukemia or lymphoma as children, two to three malignancies might have been expected in this group of 18 patients, but none were observed.

If expression of normal transcript from the 4-1-8 allele moderates the phenotype, there may be a quantitative relationship between the level of normal product and the phenotype. There is clearly some degree of variation among 4-1-8 patients. This is particularly evident in the ages at which significant gait problems become apparent (e.g., at 1 year of age in patient 52-4, but not until 12 years of age in patient 62-4 [table 1]). It is of interest that in both patients, 44-4 and 62-4, the level of expression of the abnormal 4-1-8 transcript is low (fig. 2) perhaps reflecting a higher level of normal splicing from the 4-1-8 allele. These observations may have important implications for the development of gene therapy for A-T, since it is possible that even low levels of expression of the ATM gene may have a significant beneficial effect on the clinical phenotype.

In addition to the patients associating the 137-bp insertion with a less severe phenotype, we can show that other mutations can also give rise to a similar less severe clinical and cellular appearance. It is striking that, in a second group of four patients with these features, two patients, in whom mutations have been identified, had different missense mutations, each of which may contribute to the less severe phenotype although possibly in different ways.

Several papers have published information on the spectrum of mutations in A-T patients (Savitsky et al. 1995*a;* Byrd et al. 1996; Gilad et al. 1996). There are many different mutations distributed over most or all parts of the gene. About 89% of mutations are reported to inactivate the ATM protein by truncating it and a proportion of the remainder are speculated to give rise to a less-severe phenotype (Gilad et al. 1996). The second mutation, in some patients, remains to be identified, and it is possible that a proportion may be intronic. Functionally, however, mutations may be divided into two classes. The first class includes those that result in substantial or complete loss of protein function, e.g., truncating mutations that appear to make up the majority of mutations detected to date (Savitsky et al. 1995*a;* Byrd et al. 1996; Gilad et al. 1996). The mutational mechanism in many of these cases appears to be the deletion of one or a few nucleotides because of polymerase slippage, since deletion was always seen to occur within a short stretch of mono- or dinucleotide repeats. Deletions of small numbers of amino acids at important functional positions may also have a similar effect. The second class includes mutations that have fewer detrimental effects on protein structure/function, e.g., missense mutations, or incompletely penetrant mutations such as the 4-1-8 mutation. It is clear from this that there is considerable potential for clinical heterogeneity, since different combinations of class I and II mutations may give rise to different phenotypes. We propose that a class I/I genotype results in the classical A-T phenotype. This is consistent with the proposal of Gilad et al. (1996) that homozygosity or compound heterozygosity

for null alleles is likely to give rise to the features of classical A-T. In contrast, the variant patients described here represent a class I/II genotype. The phenotype of individuals with a II/II genotype provides an interesting area for speculation. It is possible that such individuals do not have a clearly recognizable A-T phenotype but may show some of the individual features of A-T to a variable degree, e.g., radiosensitivity, genetic instability, or mild neurological or immunological features. The occurrence of a class II/II genotype might also provide at least part of the explanation for the observation of a lower than expected rate of consanguinity in A-T. The relatively low level of consanguinity suggests that the mutant gene is more widely distributed in the general population than is predicted from incidence rates.

Acknowledgments

We thank the Cancer Research Campaign, the Wellcome Trust, the U.K. A-T Society, the A-T Research and Support Trust, and the A-T Medical Research Trust for their continued support. We would also like to thank The University of Birmingham Medical School for support and Alta Bioscience, University of Birmingham, for sequencing and genotyping.

References

Brown EJ, Albers MW, Shin TB, Ichikawa K, Keith CT, Lane WS, Schreiber SL (1994) A mammalian protein targeted by G1-arresting rapamycin-receptor complex. Nature 369: 756–758

Byrd PJ, McConville CM, Cooper P, Parkhill J, Stankovic T, McGuire GM, Thick JA, et al (1996) Mutations revealed by sequencing the 5′ half of the gene for ataxia telangiectasia. Hum Mol Genet 5:145–149

Gilad S, Khosravi R, Shkedy D, Uziel T, Ziv Y, Savitsky K, Rotman G, et al (1996) Predominance of null mutations in ataxia telangiectasia. Hum Mol Genet 5 :433–439

Hari KL, Santerre A, Sekelsky JJ, McKim KS, Boyd JB, Hawley RS (1995) The mei-41 gene of D. melanogaster is a structural and functional homolog of the human ataxia telangiectasia gene. Cell 82:815–821

Hollman M, Boulter J, Maron C, Beasley L, Sullivan J, Pecht G, Heinman S (1993) Zinc potentiates agonist-induced currents at certain splice variants of the NMDA receptor. Neurone 10:943–954

Jimenez G, Yucel J, Rowley J, Subramani S (1992) The rad 3^{+}gene of *Schizosaccharomyces pombe* is involved in multiple checkpoint functions and in DNA repair. Proc Natl Acad Sci (USA) 89:4952–4956

Lelli N, Garuti R, Ghisellini M, Tiozzo R, Rolleri M, Aimale V, Ginocchio E, et al (1995) Occurrence of multiple aberrantly spliced messenger-RNAs of the LDL-receptor gene upon a donor splice-site mutation that causes familial hypercholesterolemia. J Lipid Res 36:1315–1324

Liu Q, Sommer SS (1995) Restriction endonuclease fingerprinting (REF): a sensitive method for screening mutations in long, contiguous segments of DNA. Biotechniques 18: 470–477

McConville CM, Byrd PJ, Ambrose HJ, Taylor AMR (1994) Genetic and physical mapping of the ataxia-telangiectasia locus on chromosome 11q22–q23. Int J Radiat Biol 66: S45–S56

Menon KP, Neufeld EF (1994) Evidence for degradation of mRNA encoding alpha-L-iduronidase in Hurler fibroblasts with premature termination alleles. Cell Mol Biol 12:1149–1161

Sabatini DM, Erdjument-Bromage H, Lui M, Tempst P, Snyder SH (1994) *RAFT1:* a mammalian protein that binds to FKBP12 in a rapamycin dependent fashion and is homologous to yeast TORs. Cell 78:35–43

Savitsky K, Bar-Shira A, Gilad S, Rotman G, Ziv Y, Vanagaite L, Tagle DA, et al (1995*a*) A single ataxia telangiectasia gene with a product similar to PI-3 kinase. Science 268: 1749–1753

Savitsky K, Sfez S, Tagle DA, Ziv Y, Sartiel A, Collins FS, Shiloh Y, et al (1995*b*) 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

Seaton BL, Yucel J, Sunnerhagen P, Subramani S (1992) Isolation and characterisation of the *Schizosaccharomyces pombe* rad 3 gene, involved in the DNA damage and DNA synthesis checkpoints. Gene 119:83–89

Sedgwick RP, Boder E (1991) Ataxia-Telangiectasia. In: de Jong JMBV (ed) Handbook of clinical neurology: hereditary neuropathies and spinocerebellar atrophies, vol 16. Elsevier, Amsterdam, pp 347–423

Shapiro MB, Senapathy P (1987) RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. Nucleic Acids Res 15:7155–7174

Spector BD, Filipovich AH, Perry GS, Kersey KS (1982) Epidemiology of cancer in ataxia telangiectasia In: Bridges BA, Harnden DG (eds) Ataxia telangiectasia—a cellular link between cancer, neuropathology, and immune deficiency. John Wiley, Chichester, pp 103–138

Stamm S, Zhang MQ, Marr TG, Helfman DM (1994) A sequence compilation and comparison of exons that are alternatively spliced in neurones. Nucleic Acids Res 22:1515–1526

Szabo CI, King M-C (1995) Inherited breast and ovarian cancer. Hum Mol Genet 4:1811–1817

Taylor AMR, McConville CM, Rotman G, Shiloh Y, Byrd PJ (1994) A haplotype common to intermediate radiosensitivity variants of ataxia-telangiectasia in the UK. Int J Radiat Biol 66:S35–S41

Taylor AMR, Hernandez D, McConville CM, Woods CG, Stacey M, Biggs P, Byrd PJ et al (1996*a*) Malignant disease and variations in radiosensitivity in ataxia telangiectasia patients. In: Eeles RA, Ponder BAJ, Easton DF, Horwich A (eds) Genetic predisposition to cancer. Chapman and Hall, London, pp 138–151

Taylor AMR, Metcalfe JA, Thick J, Mak Y-F (1996*b*) Lymphoma and leukaemia in ataxia telangiectasia. Blood 87:423–438

Uziel T, Savitsky K, Platzer M, Ziv Y, Helbitz T, Nehls M, Boehm T, et al (1996) Genomic organisation of the ATM gene. Genomics 33:317–320