© 2004 Nature Publishing Group http://www.nature.com/naturegenetics

Senataxin, the ortholog of a yeast RNA helicase, is mutant in ataxia-ocular apraxia 2

Maria-Céu Moreira[1], Sandra Klur[1], Mitsunori Watanabe[2], Andrea H Németh[3], Isabelle Le Ber[4], José-Carlos Moniz[5], Christine Tranchant[6], Patrick Aubourg[7], Meriem Tazir[8], Lüdger Schöls[9], Massimo Pandolfo[10], Jörg B Schulz[11], Jean Pouget[12], Patrick Calvas[13], Masami Shizuka-Ikeda[2], Mikio Shoji[2], Makoto Tanaka[2], Louise Izatt[14], Christopher E Shaw[15], Abderrahim M'Zahem[16], Eimear Dunne[3], Pascale Bomont[1], Traki Benhassine[17], Naïma Bouslam[4], Giovanni Stevanin[4], Alexis Brice[4], João Guimarães[18], Pedro Mendonça[19], Clara Barbot[20,21], Paula Coutinho[20,22], Jorge Sequeiros[20], Alexandra Dürr[4], Jean-Marie Warter[6] & Michel Koenig[1]

Ataxia-ocular apraxia 2 (AOA2) was recently identified as a new autosomal recessive ataxia. We have now identified causative mutations in 15 families, which allows us to clinically define this entity by onset between 10 and 22 years, cerebellar atrophy, axonal sensorimotor neuropathy, oculomotor apraxia and elevated alpha-fetoprotein (AFP). Ten of the fifteen mutations cause premature termination of a large DEAxQ-box helicase, the human ortholog of yeast Sen1p, involved in RNA maturation and termination.

We previously identified a 16-cM interval on chromosome 9q34 associated with an autosomal recessive adolescent-onset cerebellar ataxia segregating in two families[1,2], one with additional oculomotor apraxia[1] and the second with associated elevated serum AFP, immunoglobulins and creatine kinase levels but no oculomotor apraxia[2,3]. We identified nine additional families with ataxia linked to 9q34 by homozygosity mapping (**Supplementary Methods** online). As most affected individuals had both oculomotor apraxia and elevated AFP levels we assumed that they were affected by the same disorder, which we named AOA2 (OMIM 606002). We identified distal and proximal recombinations in families with two affected individuals (**Fig. 1a**), localizing the defective gene underlying AOA2 to a 1.1-Mb interval containing 13 genes (**Fig. 1b**) and three groups of overlapping spliced expressed-sequence tags, which we analyzed for nucleotide changes but found no mutations. We also found that the unspliced mRNA AK024331 overlaps with the *KIAA0625* cDNA and is part of a larger transcript overlapping with additional exons on the 5′ side. We obtained an open reading frame of 8,031 nucleotides and 24 exons (**Fig. 1c**), of which exon 8 was 4,177 nucleotides long. We confirmed the prediction and size of the transcript by long-range RT-PCR experiments spanning the putative exon 1 and 3′ untranslated region in human fibroblast and lymphoblastoid cell lines (data not shown) and by hybridization of a human northern blot with a probe spanning putative exons 8–24 (**Fig. 1d**). We also identified an alternative transcript that is 2.4 kb longer, resulting from a second polyadenylation site (human mRNAs AB014525 and AK022902; **Fig. 1d**).

We sequenced exons 1–18 and flanking intronic sequences in families with ataxia linked to this region and in additional individuals with either AOA or ataxia with elevated AFP levels and found 15 different disease-associated mutations in 15 families (**Table 1**). Ten of these mutations, including mutations in the two families in whom we first identified AOA2, cause truncation of the protein, indicating that this is the gene underlying AOA2. We found the nonsense mutation R1363X in three unrelated families originating from Portugal, Cabo Verde (once a Portuguese colony) and Spain, suggestive of an Iberian founder event, although recurrent C→T changes on this CpG dinucleotide cannot be formally excluded. Absence of the five missense mutations in 150 unrelated and unaffected individuals sharing the same ethnic origin as the affected individuals indicates that they are not frequent polymorphisms. Two of the missense mutations were associated with a frameshift mutation inherited from the other parent, and the remaining missense mutations were present in the homozygous state in the affected individuals. We identified four variants resulting in amino acid changes and a silent nucleotide change (**Table 1**) on the normal chromosome of healthy siblings or parents from several families, indicating that they were frequent polymorphisms.

Before our mapping, the disorders in the different families were considered to be clinically distinct entities. We can now delineate the common clinical phenotype associated with mutant senataxin, illustrating the power of defining disorders by their genetic locus and identified mutations. We considered only those families in whom we had confirmed mutations when delineating the AOA2 phenotype, as some consanguineous families with sporadic affected individuals could show

[1]IGBMC (Centre National de la Recherche Scientifique, Institut National de la Santé et de la Recherche Médicale, ULP) 67404 Illkirch, C.U. de Strasbourg, France. [2]Gunma University School of Medicine, Maebashi 371-8511, Japan. [3]Wellcome Trust Centre for Human Genetics, Oxford OX3 7BN, UK. [4]Institut National de la Santé et de la Recherche Médicale U289 and Département de Génétique, Cytogénétique et Embryologie, Hôpital de la Salpêtrière, AP-HP, 75651 Paris, France. [5]Hospital Sto. António dos Capuchos, Lisboa, Portugal. [6]Hôpitaux Universitaires de Strasbourg, 67091 Strasbourg, France. [7]Institut National de la Santé et de la Recherche Médicale U342, Hôpital Saint Vincent de Paul, 75675 Paris, France. [8]Centre Hospitalier Universitaire Mustapha, Algier 16000, Algeria. [9]Department of Neurology, St. Josef Hospital, Ruhr-University Bochum, Germany. [10]Université Libre de Bruxelles-Hôpital Erasme, 1070 Brussels, Belgium. [11]Department of General Neurology, Center of Neurology and Hertie Institute for Clinical Brain Research, University of Tübingen, 72076 Tübingen, Germany. [12]Hôpital de la Timone Adultes, 13385 Marseille, France. [13]Hôpital Purpan, 31059 Toulouse, France. [14]Department of Clinical Genetics, Guy's Hospital, London SE1 9RT, UK. [15]Guy's King's and St Thomas' School of Medicine and Institute of Psychiatry, King's College, London, SE5 8AF, UK. [16]Centre Hospitalier Universitaire Ben Badis, Constantine, Algeria.[17]Institut Pasteur d'Alger, Algeria. [18]Hospital de Egas Moniz, Lisboa, Portugal. [19]Department of Hematology, Hospital do Divino Espírito Santo, S. Miguel, Azores, Portugal. [20]UnIGENe - IBMC, ICBAS, Universidade Porto, 4150 – 180 Porto, Portugal. [21]Department of Neuropediatrics, Hospital Maria Pia, 4050 Porto, Portugal. [22]Department of Neurology, Hospital São Sebastião, 4520 Sta. Maria da Feira, Portugal. Correspondence should be addressed to M.K. (mkoenig@igbmc.u-strasbg.fr).

Published online 8 February 2004; doi:10.1038/ng1303

homozygosity at 9q34 by chance rather than by linkage. AOA2 shares several clinical features with AOA1, including gait ataxia, cerebellar atrophy, sensory-motor neuropathy (93%) and oculomotor apraxia (47%), but can be distinguished by a later onset (at 10–22 years of age in AOA2 versus 2–15 years of age in AOA1), high levels of AFP (86%, although normal laboratory values are highly variable; **Supplementary Table 1** online) and normal serum albumin levels after a long disease duration. Other features were more variable, including choreoathetosis, dystonic movements and elevated serum levels of creatine kinase. Some features of the core clinical phenotype for AOA2 are also seen in ataxia-telangiectasia and ataxia-telangiectasia–like disorder[4,5]. Because individuals with the latter conditions are predisposed to solid tumors, protein-based assays to distinguish between AOA2 and ataxia-telangiectasia and ataxia-telangiectasia–like disorder are needed.

The predicted protein encoded by the gene mutated in AOA2 is 2,677 amino acids long and contains at its C terminus a classical seven-motif domain found in the superfamily 1 of helicases[6]. In particular, it shares extensive homologies with the fungal Sen1p proteins (**Fig. 1e**), and so we named it senataxin (SETX). *Saccharomyces cerevisiae* Sen1p is involved in splicing and termination of tRNA, small nuclear RNA and small nucleolar RNA and has RNA helicase activity encoded by its C-terminal domain[7–9]. *Schizosaccharomyces pombe* has two *Sen1* genes. The first reported *S. pombe* Sen1p (Sen1p1, encoded on chromosome I) has both RNA and DNA helicase activities[9]. Of all the fungal Sen1p proteins, however, the second *S. pombe* Sen1p (Sen1p2, encoded on chromosome II) has the highest homology with senataxin over the N-terminal domain (20% identity over 466 residues). The C-terminal

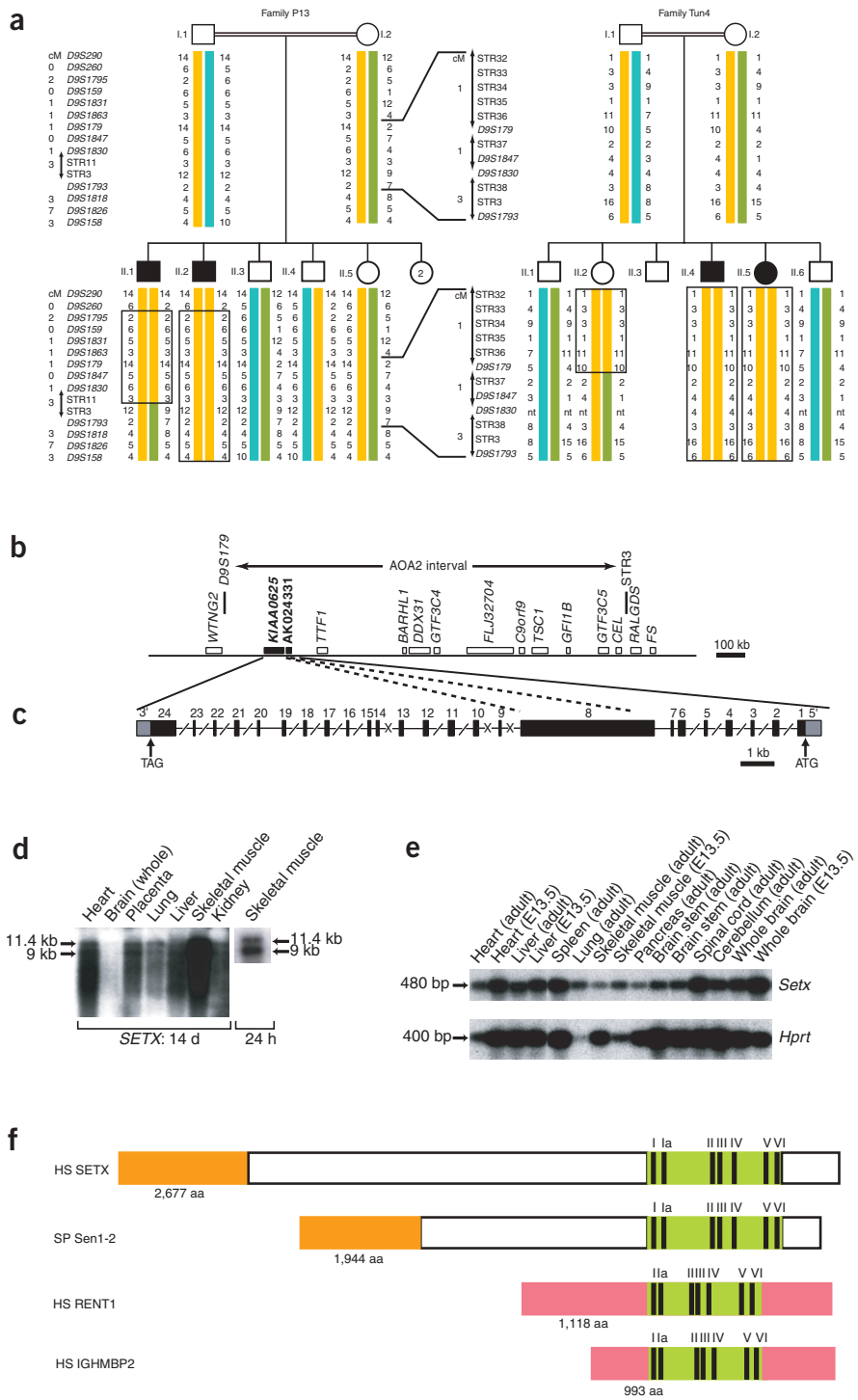

Figure 1 Genetic and transcriptional analysis of the AOA2 region. (**a**) 9q34 haplotype segregation in families P13 and Tun4. Markers are indicated on the left of each genealogical tree, in order from centromere to telomere. Distances in cM are indicated. Regions of homozygosity are boxed. Individual II.2 of family Tun4 was normal on neurological examination at 33 years of age, well beyond the age of onset for her affected brother and sister (21 years). (**b**) Transcriptional map of the critical region associated with AOA2. The *KIAA0625* cDNA and AK024331 mRNA are part of *SETX*. (**c**) Exon organization of *SETX*. Exons (boxes) and small introns (lines) are represented to scale. Introns above 1 kb or 7 kb in length are interrupted with a slash or an X, respectively. (**d**) Human multiple tissue northern blot (Clontech) analysis. Tissues are indicated above each lane. RNA from the brain and liver lanes was partially degraded. Autoradiography of the skeletal muscle lane from the same experiment after 24 h is also shown, to allow resolution of the two isoforms. Autoradiography exposure time is indicated below. (**e**) Semiquantitative RT-PCR analysis of adult and fetal (embryonic day (E) 13.5) tissues for mouse *Setx* and *Hprt* transcripts. *Setx* is ubiquitously expressed. *Hprt* transcript analysis of the same reverse transcription samples serves as a control for the amount of analyzed tissue. (**f**) Domain organization of senataxin (SETX) and related proteins. The N-terminal domain shared between the SETX and Sen1 proteins is indicated in orange. The helicase domain is indicated in green. The position of the seven helicase motifs (I–VI) is indicated on top. The intermediate domain (white) is weakly conserved between mammalian SETX and fungal Sen1 proteins. Domains specific to the RENT1 or to the IGHMBP2 proteins have a distinct color code. HS, *Homo sapiens*; SP, *S. pombe*. Total amino acid (aa) length is indicated below the protein boxes.

© 2004 Nature Publishing Group http://www.nature.com/naturegenetics

Table 1 Mutations and polymorphisms in *SETX*

Family (geographic origin)	Nucleotide change (exon)	Amino acid change	Mutation status
Mutations			
Alg5 (Algeria)	2602C→T (8)	Q868X	Homozygous
Alg11 (Algeria)	915G→T (6)	W305C	Homozygous
Can7 (Canada)	5070insT (8)	fs after I1690	Homozygous
F30 (Vietnam)	2622–2625delAGTT (8)	fs after L874	Homozygous
F77 (France)	994C→T (6)	R332W	Compound
	2966–2970delGGAAA (8)	fs after Q988	Heterozygous
F88 (France)	5264delC (8)	fs after N1754	Heterozygous
P13 (Cabo Verde), F93 (Portugal) and G10 (Spain)	4087C→T (8)	R1363X	Homozygous
F98 (Turkey)	5249insT (8)	fs after L1750	Homozygous
J7 (Japan)	4321C→T (8)	Q1441X	Homozygous
J8 (Japan)	6638C→T (18)	P2213L	Homozygous
Pak1 (Pakistan)	2332C→T (8)	R788X	Homozygous
Tun4 (Tunisia)	1238C→T (8)	P413L	Homozygous
UK1 (United Kingdom)	879delT (6)	fsC292	Compound
	5267T→C (8)	F1756S	Heterozygous
Polymorphisms			
F89, J10, Tur1	3147C→T (8)	H1049H	Homozygous
F88, F89, F90, F91, G10, J8, J10, P13, Tun4, Tur1	3455G→T (8)	C1152F	Homozygous
F90, F91, G10, P13, Tun4	3576T→G (8)	D1192E	Homozygous
F88, F90, F91, G10, P13, Tun4	3754G→A (8)	G1252R	Homozygous
Alg5, F11, F77, F89, F91, F102, J7, J8, J10, Pak1, P13, Tur1, UK1	4156G→A (8)	V1386I	Homozygous

fs, frameshift.

domain of senataxin and the Sen1p proteins shares significant similarity with two other members of the DExxQ-box family of helicases (**Fig. 1f**): RENT1/Upf1, involved in nonsense mediated RNA decay[10], and IGHMBP2, defective in spinal muscular atrophy with respiratory distress[11] (OMIM 604320), a human disorder of motor neurons, and in mouse neuromuscular degeneration[12]. Upf1 proteins have RNA helicase activity, but IGHMBP2 was initially identified as a DNA binding protein with transcriptional transactivating properties[13]. It is therefore possible that, like *S. pombe* Sen1p1, senataxin has both RNA and DNA helicase activities and that senataxin acts in a DNA repair pathway, like several other proteins defective in autosomal recessive cerebellar ataxias, as in ataxia-telangiectasia[4], AOA1 (ref. 14), ataxia-telangiectasia–like disorder[5] and spinocerebellar ataxia with peripheral neuropathy 1 (ref. 15). Alternatively, the results also suggest that senataxin might be a nuclear RNA helicase with a role in the splicing machinery and that the molecular pathology of AOA2 may share features with spinal muscular atrophy and spinal muscular atrophy with respiratory distress. Our results add to the increasing evidence to suggest that both DNA repair and RNA splicing are key factors in several neurodegenerative disorders, including the newly identified AOA2, and further work may elucidate the role of these mechanisms in neuronal integrity and neurodegeneration.

GenBank accession numbers. Human *SETX*, AY362728; mouse *Setx*, BK001523. Human cDNA and mRNAs: *KIAA0625*, NM_015046, 31543019; AK024331, 10436690; AK022902, 10434561; AB014525, 3327063. Mouse mRNAs: AK044730, 26336742; AK048354, 26339275. Human expressed-sequence tags: BC032622, 22749753; CB162163, 28148289; AA578438, 2356622; AW812833, 7905827; AI216401, 3785442. Mouse expressed-sequence tags: BU503417, 22809650; BY718091, 27131208. Sen1 proteins: *Neurospora crassa*, AL442164, 16945408; *S. pombe* (Sen1p2), NC_003423, 19112847.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

This manuscript is dedicated to the memory of J.-M. Warter, who died on 19 September 2003. We thank the affected individuals and families for their collaboration; D. Simon, A. Buj-Bello, C. Alves, L. Pereira, A. Amorim, D. Grid and A. Hamri for sharing biological material; J.-L. Mandel for his support, encouragement and discussions; and E. Troesch, F. Ruffenach, I. Colas, S. Vicaire, M. Amari and K. Mizushima for technical help. Genetic studies were supported by funds from the Fundação para a Ciência e a Tecnologia (Portuguese Ministry of Science) and from the Portuguese Ministry of Health, the Institut National de la Santé et de la Recherche Médicale, the Centre National de la Recherche Scientifique, the Hôpitaux Universitaires de Strasbourg (PHRC regional to J.M.W.) and the GIS-Maladie Rares (to A.D.). M.C.M. has postgraduate fellowship from Fundação para a Ciência e a Tecnologia (Portuguese Ministry of Science). A.H.N. and E.D. were supported by the US National Ataxia Foundation. S.K. and P.B. were supported by fellowships from the Association Française contre l'Ataxie de Friedreich.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Received 17 September 2003; accepted 21 January 2004
Published online at http://www.nature.com/naturegenetics/

1. Németh, A. *et al. Am. J. Hum. Genet.* **67**, 1320–1326 (2000).
2. Bomont, P. *et al. Eur. J. Hum. Genet.* **8**, 986–990 (2000).
3. Watanabe, M. *et al. Ann. Neurol.* **44**, 265–269 (1998).
4. Shiloh, Y. *Nat. Rev. Cancer* **3**, 155–168 (2003).
5. Stewart, G.S. *et al. Cell* **99**, 577–587 (1999).
6. Gorbalenya, A.E. & Koonin, E.V. *Curr. Opin. Struct. Biol.* **3**, 419–429 (1993).
7. Ursic, D., Himmel, K.L., Gurley, K.A., Webb, F. & Culbertson, M.R. *Nucleic Acids Res.* **25**, 4778–4785 (1997).
8. Rasmussen, T.P. & Culbertson, M.R. *Mol. Cell. Biol.* **18**, 6885–6896 (1998).
9. Kim, H.D., Choe, J. & Seo, Y.S. *Biochemistry* **38**, 14697–14710 (1999).
10. Wang, W., Czaplinski, K., Rao, Y. & Teltz, S.W. *EMBO J.* **20**, 880–890 (2001).
11. Grohmann, K. *et al. Nat. Genet.* **29**, 75–77 (2001).
12. Cox, G.A., Mahaffey, C.L. & Frankel, W.N. *Neuron* **21**, 1327–1337 (1998).
13. Mizuta, T.R., Fukita, Y., Miyoshi, T., Shimizu, A. & Honjo, T. *Nucleic Acids Res.* **21**, 1761–1766 (1993).
14. Moreira, M.C. *et al. Nat. Genet.* **29**, 189–193 (2001).
15. Takashima, H. *et al. Nat. Genet.* **32**, 267–272 (2002).

© 2004 Nature Publishing Group http://www.nature.com/naturegenetics