Genomics **57,** 380–388 (1999)
Article ID geno.1998.5739, available online at http://www.idealibrary.com on IDEAL®

TOM1 Genes Map to Human Chromosome 22q13.1 and Mouse Chromosome 8C1 and Encode Proteins Similar to the Endosomal Proteins HGS and STAM

Eyal Seroussi,* Darek Kedra,* Maria Kost-Alimova,†,‡ Ann-Christin Sandberg-Nordqvist,* Ingegerd Fransson,* Joannes F. M. Jacobs,* Ying Fu,§ Hua-Qin Pan,§ Bruce A. Roe,§ Stephan Imreh,† and Jan P. Dumanski*,[1]

*Department of Molecular Medicine, CMM Building, L8:00, Karolinska Hospital, S-171 76 Stockholm, Sweden; †Microbiology and Tumor Biology Center (MTC), Karolinska Institutet, S-171 77 Stockholm, Sweden; ‡Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Vavilov Street 32, 117984 Moscow, Russia; and §Department of Chemistry and Biochemistry, University of Oklahoma, Norman, Oklahoma 73019

Received September 24, 1998; accepted December 28, 1998

The avian *tom1* (target of myb 1) gene has been previously characterized from v-myb-transformed cells. We report here cloning of the human and mouse *tom1* orthologs. Both genes are expressed ubiquitously, with the highest levels in skeletal muscle, brain, and intestines, as assessed by Northern blot and mRNA *in situ* hybridization. The N-terminal domain of the TOM1 protein shares similarity with HGS (hepatocyte growth factor-regulated tyrosine kinase substrate) and STAM (signal-transducing adaptor molecule), which are associated with vesicular trafficking at the endosome. A putative coiled-coil domain was also detected in the central part of the TOM1 protein. This domain structure suggests that *TOM1* is another member of a family of genes implicated in the trafficking regulation of growth-factor-receptor complexes that are destined for degradation in the lysosome. We also show that a human paralog of *TOM1* (*TOM1*-like gene 1) exists. Furthermore, we provide a transcription map over a 190-kb contig of the *TOM1* region. This map includes its distal neighbors *HMOX1* and *MCM5* and two proximal novel genes, one of which is a HMG-box-containing gene (HMG2L1), and the other of unknown function. Using a genomic PAC clone, we demonstrate that the mouse *Tom1* and *Hmox1* genes are part of an as yet undescribed syntenic group between mouse chromosome 8C1 and human chromosome 22q13.1.

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under Accession Nos. AJ006973 (*Homo sapiens* mRNA for TOM1 protein); AJ006972 (*Mus musculus* mRNA for Tom1 protein); AC005290 (*M. musculus* Chromosome 8 PAC Clone 411O19); AJ010069 (Human high mobility group box containing gene HMG2L1); AJ010070 (Human partial transcript encompassing THC211630-gene); and AJ010071 (Human *TOM1* like gene 1, TOM1L1).

[1] To whom correspondence should be addressed. Telephone: +46-8-517 73922 or 72229. Fax: +46-8-517 73909. E-mail:Jan.Dumanski@cmm.ki.se.

INTRODUCTION

The retroviral oncogene v-*myb* of avian myeloblastosis virus (AMV) and avian leukemia virus E26 is a mutated and truncated version of the *c-myb* proto-oncogene, which encodes a transcription factor, v-myb (Graf 1992). A recent differential display study of *v-myb* oncogene-transformed avian myelomonocytic cells revealed a gene designated *tom1* (for target of myb 1) (Burk *et al.*, 1997), which was a specific target of the *v-myb* oncogene. This chicken *tom1* gene has two promoters. The v-myb protein is acting via one of these and induces a shorter *tom1A* transcript form. *tom1A* is capable of encoding a truncated version of the tom1 protein, composed of 147 amino acids, compared to the 515 amino-acid-protein encoded by the longer and ubiquitously expressed *tom1B* transcript isoform. This deregulated expression of the chicken *tom1* gene has been suggested to interfere with the normal mechanism controlling proliferation and/or differentiation of myelomonocytic cells (Burk *et al.*, 1997).

Meningioma, a common tumor of the meninges covering the central nervous system, is one of numerous tumor forms that frequently display deletions on chromosome 22 (Zang and Singer 1967; Dumanski *et al.*, 1987; Seizinger *et al.*, 1987). Meningioma has been thoroughly studied with regard to the frequency and extent of tumor-specific deletions (Ruttledge *et al.*, 1994b; Lekanne Deprez *et al.*, 1995). Our previous LOH (loss of heterozygozity) studies revealed several regions on 22q likely to harbor tumor suppressor genes. One of the regions was defined by a combination of two interstitial deletions, a homozygous deletion in one tumor and a heterozygous deletion in another (Ruttledge *et al.*, 1994b), and delineated a segment of >1 Mb. These findings prompted us to investigate in detail this chromosomal region with regard to candidate genes. We detected a human ortholog of the chicken

0888-7543/99 $30.00

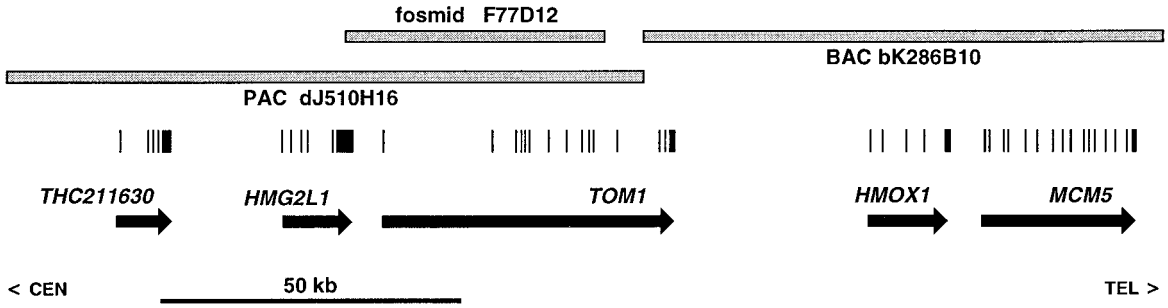

FIG. 1. Genomic structure of the human *TOM1* gene from 22q13.1. Schematic delineation of the position and transcriptional orientation of genes located in the 190-kb contig formed by PAC dJ510H16 (Accession No. AL008635), BAC bK286B10 (Accession No. Z82244), and fosmid F77D12 (Accession No. Z82097). The sequence data for these clones were produced by the Human Chromosome 22 Sequencing Group at the Sanger Centre (Hinxton, UK). The gene abbreviations are as follows: THC211630, a gene of unknown function; high mobility group box containing gene, *HMG2L1;* human ortholog of chicken target of myb 1 gene, *TOM1;* human heme oxygenase 1 gene, *HMOX1* (Kutty *et al.,* 1994); human replication protein 64, *MCM5* (Paul *et al.,* 1996). The exons for each of the genes are marked as black boxes above the arrows that indicate their transcriptional orientation. "CEN" and "TEL" indicate the positions of the centromere and telomere of 22q, respectively.

tom1 in the region. In this study we characterized the human *TOM1* and mouse *tom1* genes. We found that the N-terminal parts of the human, mouse and chicken TOM1 proteins show similarity to the N-terminal domain of the HGS, STAM and VPS27 proteins. We also demonstrate that the human *TOM1* gene and its mouse ortholog belong to an as yet undescribed syntenic group between human chromosome 22q13.1 and mouse chromosome 8C1.

MATERIALS AND METHODS

Sequencing and informatics procedures were essentially carried out as outlined in Peyrard *et al.,* 1999. Fluorescent SSCP analysis (Orita *et al.,* 1989a; Orita *et al.,* 1989b) was preformed on ABI377 instrument as described previously (Ellison *et al.,* 1993). Fluorescent in situ hybridization (FISH) was done as described by Fedorova *et al.,* 1997. The animal experiments and *in situ* hybridization histochemistry procedures were performed according to published protocols (Deschenes *et al.,* 1996; Sandberg-Nordqvist *et al.,* 1996).

RESULTS AND DISCUSSION

Cloning the Human TOM1 Gene

The chromosomal region indicated by tumor deletions and analyzed in this study is located within the large segment of the long arm of human chromosome 22, which is being systematically sequenced by the Sanger Centre (Hinxton, UK). We imported the newly generated genomic sequence (fosmid F77D12, Accession No. Z82097; PAC dJ510H16, Accession No. AL008635; and BAC bK286B10, Accession No. Z822244), joined them into a 190.3-kb contig (Fig. 1), and scanned the contig for the presence of active genes. We performed masking of the repetitive elements in the genomic sequence followed by *BlastN* analysis, which revealed that the region contains two previously cloned and mapped genes: human heme oxygenase (*HMOX1*) (Kutty *et al.,* 1994) and human replication protein (*MCM5*) (Paul *et al.,* 1996). The latter was originally characterized in yeast as the minichromosome maintenance protein Cdc46/Mcm5 (Chen *et al.,* 1992). Two additional genes were identified and partially assembled. The first was a gene containing HMG (high mobility group) box (Accession No. AJ010069), similar (amino acid identity in 30 of 76, 39%; and amino acid similarity in 45 of 76, 59%) to that of the bovine nonhistone chromosomal protein HMG-2 (Accession No. B61611). The second gene was of unknown function (Accession No. AJ010069), encompassing a group of ESTs previously assembled by the Tentative Human Consensus Effort (THC211630).

We also detected a group of 64 human ESTs that matched the human genomic sequence, and we assembled these by importing the sequence with corresponding trace files. We obtained the missing sequence data by ordering and resequencing of EST clones zr55f12 and yx98d03. The full-length cDNA sequence consisted of 2310 bp (Accession No. AJ006973). The longest open reading frame of this cDNA was capable of encoding a protein of 492 amino acids with a predicted molecular mass of 53.8 kDa. The *BlastX* search of protein databases with our cDNA revealed that it encodes a highly conserved orthologous protein (76% identity, 84% similarity) of avian *tom1* (Accession No. Y08741). Chicken *tom1* has two promoters that give rise to two transcript isoforms: the shorter (2.1 kb) *tom1A* and the longer (3 kb) *tom1B*. The latter encodes a protein of 515 amino acids (Burk *et al.,* 1997). The longest open reading frame of the human cDNA and the predicted start of translation were in agreement with those of the avian *tom1B* transcript (Fig. 2A). Another feature indicating the correct prediction of the translation start point is that a Kozak consensus for ribosome binding (9 of 13 bases) occurs in the vicinity of the first methionine codon. Comparison of the human *TOM1* cDNA with the genomic contig revealed that the gene consists of 15 exons. The genomic structure of *TOM1* was typical of other human genes (Fig. 1) with a large first intron (17.9 kb) and all splicing sites in agreement with the consensus sequence of donor (GT) and acceptor (AG). Finally, an atypical polyadenylation signal, AAT-

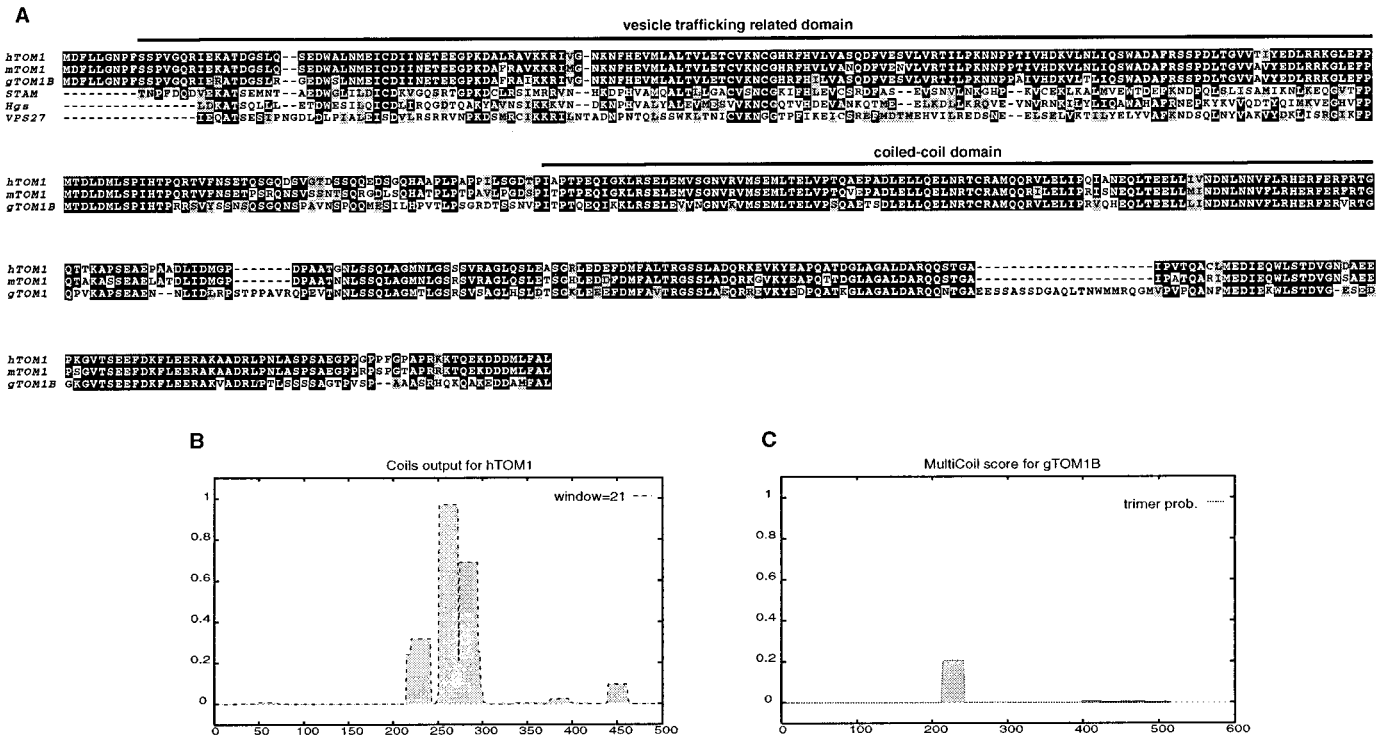

FIG. 2. (**A**) An alignment of predicted amino acid sequences of three *TOM1* genes from *Homo sapiens (hTOM1), Mus musculus (mTom1),* and *Gallus gallus (gtom1B)* and N-terminal portions of proteins implicated in the regulation of vesicular traffic: human signal-transducing adaptor molecule (STAM); human hepatocyte growth factor-regulated tyrosine kinase substrate (HGS); and yeast vacuolar-protein sorting-associated protein (VPS27). Identity and similarity between the amino acid sequences are indicated by black and gray boxes, respectively. White boxes indicate nonconservative amino acid changes between the proteins. Dashes indicate gaps introduced by the alignment program. Two protein domains of the human TOM1 protein predicted by computer-assisted sequence analysis are marked by horizontal bars above the human TOM1 protein sequence: a domain related to STAM, HGS, and VPS27 proteins (S^{10}-P^{152}) and a potential coiled-coil domain (I^{200}-G^{307}). (**B**) The sequence of human TOM1 protein was searched for coiled-coil structures with the *coils* 2.1 program using MTIDK matrix with weights and scanning windows of 21 amino acids as previously described (Lupas *et al.,* 1991). The putative coiled-coil sequences were detected at amino acid positions I^{200} through G^{307}, with a maximal potential ($P = 0.97$) to form a coiled-coil structure. (**C**) The amino acid sequences of avian tom1B protein were searched for three-stranded coiled-coil sequence with the *multicoils* program (Wolf *et al.,* 1997). The putative three-stranded coiled-coil sequences were detected between amino acid position T^{213} and V^{242}, with a maximal score (0.206) to form a coiled-coil structure. It should be noted that similar coiled-coil predictions, using *coils* and *multicoils* programs, were obtained for all TOM1 orthologs. The position in the protein sequences of human TOM1 and avian tom1B proteins is shown on the *X*-axis. On the *Y*-axis the significance scores of predictions are displayed.

TAAA, was detected in the 3′ end of the human cDNA sequence.

As shown in Fig. 2A, the predicted human TOM1 protein is smaller than the avian protein due to four deletions. The largest of these deletions consists of 21 amino acids, close to the C-terminal end of the human protein. This difference corresponds to the exon–intron boundaries of the human gene, between exons 12 and 13. Therefore, we searched in the human genomic sequence of *TOM1* for sequence strings that might encode these (or similar) 21 amino acids. We did not detect such sequence or any transcript isoforms containing this presumed extra exon. Thus, it is unlikely that this exon exists in *Homo sapiens* and *Mus musculus* (see below). It should be pointed out that the avian *tom1A* transcript also lacks this exon (Burk *et al.,* 1997).

Promoter Structure of the Human TOM1 Gene

The genomic sequence at the 5′ end of *TOM1* contains a CpG island typical of housekeeping genes. This island, predicted by the *Xgrail* computer program (402 bp with CpG score 0.813, 71.1% G+C), stretches over the promoter region and 50 bp into the first intron. The *tssw* software predicted the transcription start site, which was in agreement with our cDNA sequence (significance score 12.9). This program detects putative transcription factor binding sites and includes multiple SP1 binding sites, cAMP responsive elements (Hirai and Yaniv 1989; Monla *et al.,* 1995), and an 8-bp binding site of the early growth response gene (*EGR4*) inducible element in resting cells upon growth stimulation (Zipfel *et al.,* 1997). However, neither the Myb binding site nor any other putative promoter that could produce an alternative transcript form was present in the genomic sequence of human *TOM1.* Such a shorter transcript would be a counterpart of the avian *tom1A* isoform observed in transformed chicken myelomonocytic cells.

Characterization of the Mouse Tom1 Gene on cDNA and Genomic Level

Comparison of the human *TOM1* cDNA sequence with the mouse ESTs, deposited in *dbEST,* revealed

TABLE 1

Oligonucleotides and PCR Primers

No.	Lab ref.	Primer sequence (5′-3′)	Modification
1	37U	TTGGTGGCAGCGGCGGTAG	
2	949U	CGCCATGAACGGTTTGAA	
3	1795L	TCCTGGGGCCTCCACTT	
4	2195L	GGTGCCCGGCCAGAAGGAA	
5	ex1L	AGGCTTCCCAGGGGACCGA	HEX
6	ex2U	CAGGACCCCCTCATGACTTT	TET
7	ex2L	TGGGATGGCCGTCTCTG	HEX
8	ex3U	TTCCTAAGCCCACCCTTTT	6-FAM
9	ex3L	CCTGCCACAGGACAGACAG	HEX
10	ex4U	GTGCTCGACGGCACCTCTC	TET
11	ex4L	GCCCTGTCCTGGCACTCAC	6-FAM
12	ex5U	TGGTTACCGGCTGTGTCC	TET
13	ex5L	CGGTACGGCAGTTCTCTCA	HEX
14	ex6U	GCCCTCACTGATCCTGTTT	HEX
15	ex6L	CCCCAGGCTCGTTTACC	TET
16	ex7U.B	CACCACGATCAACTGTGTG	6-FAM
17	ex7L.B	AGCTGGGCAGGAGTCATC	HEX
18	ex8U	TGACTGTGGTTGCCTCACT	TET
19	ex8L	GGAGGTGCTGGGGCTAC	HEX
20	ex9U	GGGGCCTGTCTCAGTCTG	6-FAM
21	ex9L	GGAGGTGAGAATGGCACTC	HEX
22	ex10U	GGGGCAGAAAGGAACGAAG	TET
23	ex10L	GGAGCCCTCCAAACCACAC	HEX
24	ex11U	TCACGTGGCCTTTCTGTCC	6-FAM
25	ex11L	AGGGCCAGGCCACTCACT	HEX
26	ex13U	ACCTAGGGCACTGCCTTAC	TET
27	ex13L	GGGTCATTTCTTGGTGAGA	HEX
28	41U	CGCCGGTTGCTGTCACCT	
29	1978L	TTTTTCGGGCCCCACTCTG	
30	1520U	CCCCTCGGAGGAAGACCCA	
31	ex3u235	AAAGATGCCTTCCGAGCAG	
32	ex41420	GTCATGCACAATGGTGGGT	
33	ex4u398	CCATCCTGCCGAAGAACAA	
34	ex51473	AGGTGAGCTGCGGAACG	
35	ex61646	GGTCTCCACGCTGACTGGT	
36	ex5u567	GCTGTCGCCCATCCAC	
37	ex2u176	AGGACTGGGCCCTCAACAT	
38	ex31233	GCTCGGAAGGCATCTTTGG	
39	ex1u48	GACTTTCTCCTGGGAAATCCGTTCAGCTCTCCGGTGGGACAACGGATC	^{35}S-dATP
40	ex1148	GATCCGTTGTCCCACCGGAGAGCTGAACGGATTTCCCAGGAGAAAGTC	^{35}S-dATP

Note. Primer 1 and primers 5 through 27 were used for the F-SSCP analysis. These primers, except primer 1, were all labeled at the 5′ end with a fluorescent dye as indicated in the "modification" column. These F-SSCP primers were used for PCR in pairs as follows: 1 + 5, 6 + 7, 8 + 9, etc. Oligonucleotides 39 and 40 were used for *in situ* hybridization histochemistry and were labeled at the 3′ end by ^{35}S-dATP.

16 clones that belonged to the mouse ortholog, the *Tom1* gene. These ESTs initially were grouped in two contigs that allowed us to design PCR primers (primers 28 and 29, Table 1) for PCR amplification from the mouse brain cDNA library. Sequencing of the 1.9-kb PCR product together with the sequences of ESTs resulted in one contig corresponding to the full-length cDNA (2253 bp) of mouse *Tom1* (Accession No. AJ006972). The longest open reading frame is capable of encoding a protein of 492 amino acids that displays 89% identity and 93% similarity to the human TOM1 protein (Fig. 2A). Comparison of human and mouse cDNA sequences at the nucleotide level within the coding region revealed 87% identity. Generally, the noncoding 5′ and 3′ regions of human and mouse cDNAs were poorly conserved between the species. However, three regions were preserved pointing to their possible functional importance. These regions were a 5′ end sequence (50 of 62 bases conserved, 81%) located prior to the Kozak consensus sequence for ribosome binding; a region of 26 bp following the stop codon (23 of 26 bases conserved, 88%); and 79-bp region at the 3′ end of the cDNAs, which included a polyadenylation signal AATAAA sequence (71 of 79 bases conserved, 89%).

Using a 3′-end mouse cDNA probe, derived from PCR amplification of the mouse brain cDNA library (primers 29 and 30, 460-bp PCR product), we screened a 129Sv-strain mouse genomic PAC library (RPCI-21, Roswell Park Cancer Institute, Buffalo, NY), and three clones (411O19, 530G11, and 530H11) were detected. The PAC 411O19 was sequenced using the random shotgun method (Accession No. AC005290). This PAC spans the majority of

mouse *Tom1* from its 3′-end, including exons 2–15, and also contains the mouse ortholog of human *HMOX1*. Thus the organization of the *Tom1/Hmox1* locus in mouse is similar to the human counterpart (Fig. 1). The position of exon–intron borders in mouse *Tom1* was conserved between the human and the mouse genes, and these were annotated in the EMBL/GenBank submission file describing the mouse *Tom1* cDNA sequence (Accession No. AJ006972).

Expression of TOM1 Genes

A Northern blot hybridization (Fig. 3E) using a PCR amplified fragment of human *TOM1* cDNA (primers 1 and 4, 2.2-kb product, Table 1) revealed a single, ubiquitously expressed 2.3-kb message. Variable levels of expression were noticed, with the highest expression in skeletal muscle, placenta, heart, and liver. These results generally are in agreement with the relative abundance of ESTs from various cDNA libraries deposited in *dbEST*, with the exception of skeletal muscle. Although skeletal muscle showed the strongest signal among the tissues tested on human Northern blot, no human ESTs from this tissue could be retrieved from *dbEST*. Two mouse ESTs from a cell line related to muscle cells displayed exon skipping (AA592451, missing exon 2; AA656929, missing exon 3). No consistently shorter splicing form was detected on the human Northern blot or among the numerous human and rodent ESTs. It should be noted that the expression of avian *tom1* was most intense in skeletal muscle and also displayed a second *tom1A* transcript isoform (Burk *et al.*, 1997).

To evaluate further the expression pattern of *Tom1*, oligonucleotide probes specific for exon 1 (oligos 39 and 40, Table 1) were hybridized *in situ* to mouse embryo and adult brain tissue sections (Figs. 4A–4D). As is typical for ubiquitously expressed genes, positive hybridization signals were distributed over the whole embryos and brain parenchyma without striking differences in levels of gene expression. The highest expression was, however, noticed in the intestines (Fig. 3A). In the adult brain at the mesencephalic level, a signal was prevalent throughout the section (Fig. 3B). An intense signal was detected in the hippocampal formation and medial lemniscus. In cerebellum, strong labeling was detected in Purkinje cells and granular layers (data not shown).

As mentioned above, human *TOM1* is located in the region indicated by tumor-specific deletions and may be considered a candidate tumor suppressor gene. Therefore 12 of 15 exons present in human *TOM1* (Table 1) were tested by fluorescence SSCP in a representative series of 72 meningiomas in an attempt to determine whether this gene is affected by any point mutations within its exons or the immediately flanking intronic sequences. These tumors were selected from 170 tumors previously studied with regard to chromosome 22 genotype and mutations in the *NF2* tumor suppressor gene (Ruttledge *et al.*, 1994a; Ruttledge *et al.*, 1994b). However, no point mutations or other polymorphisms in *TOM1* could be detected within the sequences analyzed.

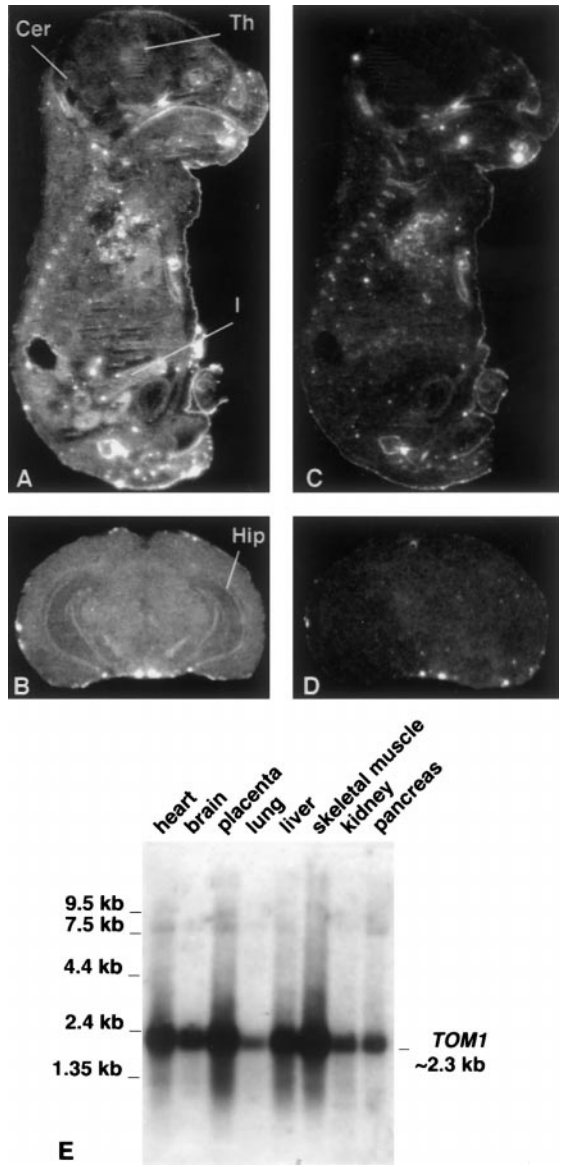

FIG. 3. The expression pattern of the *TOM1* genes in human and mouse. Dark-field autoradiograms illustrate the gene expression in mouse embryo (**A**) and in adult mouse brain (**C**) using mRNA *in situ* hybridization. A 48-mer antisense probe corresponding to first exon of the *Tom1* cDNA was hybridized to a sagittal section of a E17.5 mouse embryo (**A**) and to a coronal section from mouse adult brain at the mesencephalic level (**C**). Note the ubiquitous pattern of gene expression. The strongest signal was detected in intestines (I), which can be seen as transverse and longitudinal sections through intestines at multiple sites. A high level of expression could also be noted in the thalamic area (Th), cerebellum (Ceb), hippocampal formation (Hip) and medial lemniscus. As a negative control, the sense probe was hybridized to adjacent sections (**B** and **D**). (**E**) Ubiquitous expression of the *TOM1* gene in human tissues as an approximately 2.3-kb transcript. The entire cDNA was used as probe on Northern blot containing poly(A)$^+$ selected mRNA from eight human adult tissues.

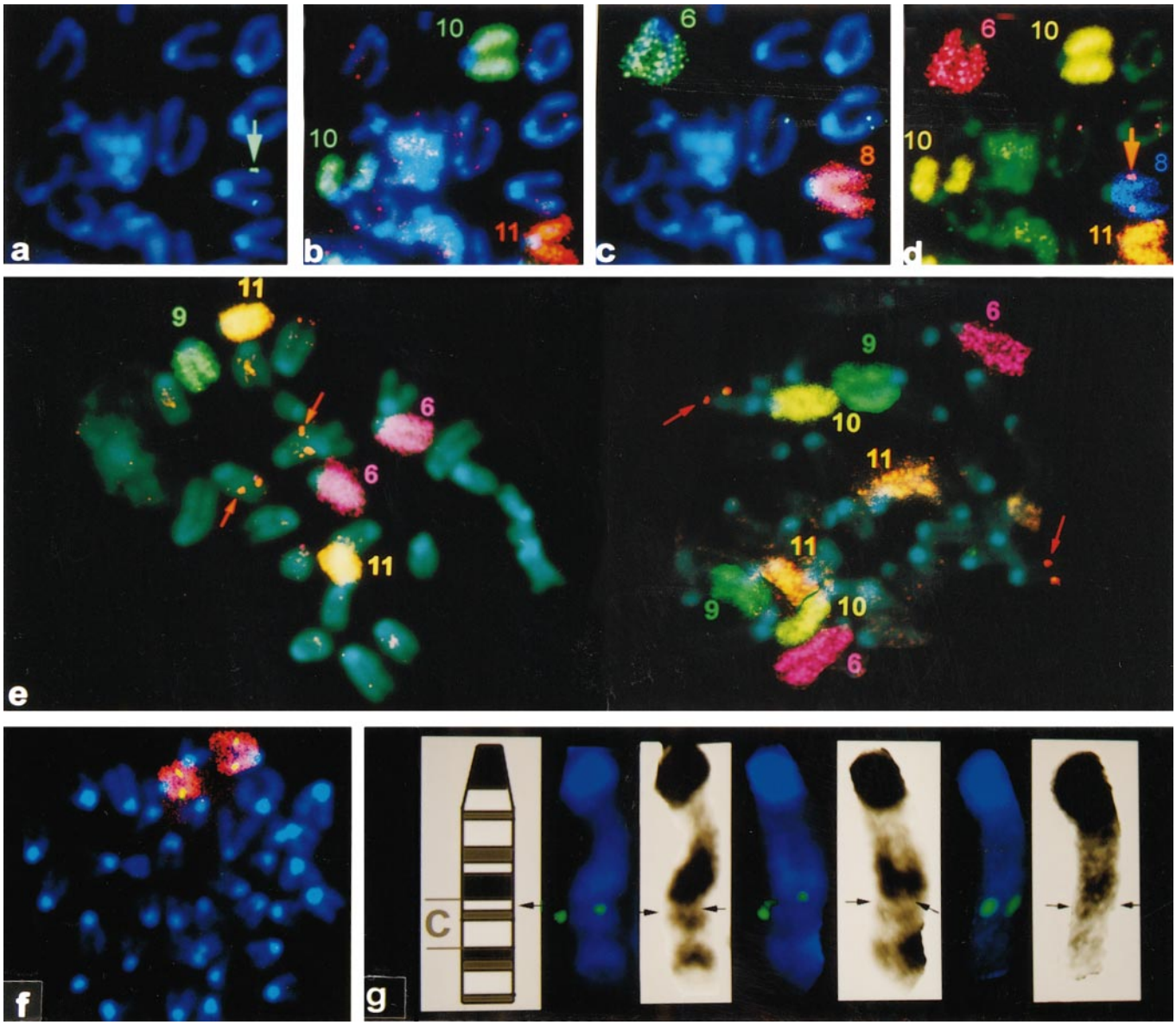

FIG. 4. Fluorescence *in situ* hybridization of PAC 411O19, containing the mouse *Tom1* and *Hmox1* genes, to mouse metaphase spreads. Chromosomes are stained with DAPI. Mouse chromosome-specific painting probes are applied for chromosome identification. (**a, b, c,** and **d**) Sequential hybridizations of PAC probe and painting probes specific for mouse chromosomes 6, 8, 10, and 11 to the same metaphase spread. (**a**) PAC 411O19 labeled with digoxigenin and detected with antidigoxigenin-FITC hybridized to mouse metaphase spreads. Arrow indicates hybridization to mouse chromosome 8. (**b**) Hybridization of FITC-labeled chromosome 10-specific and Cy3-labeled chromosome 11-specific painting probes. (**c**) Hybridization of FITC-labeled chromosome 6-specific and Cy3-labeled chromosome 8-specific painting probes. (**d**) The final result, presented as a combined pseudo-color computer image, displaying painting probes for mouse chromosomes 6, 8, 10, and 11 as well as the PAC 411O19 probe (arrow). (**e**) A pseudo-color computer image of two additional mouse metaphases after sequential hybridizations of PAC probe (arrows) and painting probes for chromosomes 6, 9, 10, and 11. (**f**) Localization of PAC 411O19 on mouse chromosome 8, which was identified by Cy3-labeled chromosome 8-specific painting probe. (**g**) PAC 411O19 was localized at chromosome 8C1 (arrows), using an inverted DAPI banding pattern.

The Predicted Domain Structure Of The Human And Mouse TOM1 Genes Provides Clues To Their Function

When the predicted amino acid sequence of *TOM1* was used in a *BlastP* database search, a family of proteins sharing a sequence similarity between their N-terminal domains was revealed. Three of these proteins have previously been characterized. These proteins are human STAM (signal-transducing adaptor molecule) protein, (Accession No. U43899); human HGS (hepatocyte growth factor-regulated tyrosine kinase substrate) protein (Accession No. U43895); and yeast VPS27 protein (SwissProt P40343). This N-terminal domain of the human TOM1 and mouse Tom1 proteins was the only part sharing similarity between them and the STAM/HGS/VPS27 proteins, which are implicated in vesicular traffic and are localized to the early endosome (Asao *et al.*, 1997; Komada *et al.*, 1997).

Therefore, this domain of the TOM1/Tom1 proteins, which displays similarity to the STAM/HGS/VPS27 proteins, was tentatively named as a "vesicle trafficking related domain" (TOM1 amino acids S^{10}-P^{152}, Fig. 2A). The notion that proteins containing this domain are vesicle trafficking regulators was further strengthened by recent reports. A splice variant of HGS, Hgs-2, was shown to be associated with SNAP25 and is thought to modulate vesicular transport (Bean *et al.*, 1997). Moreover, Lohi *et al.* (1998) suggested that the avian ortholog of STAM (60% amino acid identity, 71% similarity, which they named EAST, EGF-receptor associated protein with SH3 and TAM domains) is involved in the early stages of endocytosis and the formation of clathrin-coated vesicles. It is noteworthy that the control of vesicular transport of the activated growth factor receptor complex from early endosomes to lysosomes for degradation was recently suggested as a possible mechanism underlying the suppression of cell growth (Komada *et al.*, 1997). The HGS and STAM proteins are thought to be implicated in this process (Komada *et al.*, 1997; Asao *et al.*, 1997). HGS and STAM interact with each other to form a protein complex, presumably via coiled-coil domains present in both proteins.

Using *coils* and *multicoil* programs we predicted coiled-coil protein domains in human, mouse, and chicken tom1 proteins with a high score (human TOM1, amino acids I^{200}-G^{307}; Figs. 2A and B). The output of both programs was very similar for all three proteins. The coiled-coil domains can mediate protein–protein interactions via formation of double- or triple-helixes. The potential to form a triple-helix was best exemplified by avian *tom1B* with the *multicoil* program (Fig. 2C). Another *BlastP*-derived finding that strengthens the possible association of *TOM1/Tom1* genes with function in the intracellular vesicular transport related to the endosomal compartment was also observed. Here, the two N-terminal domains of TOM1 protein (Fig. 2) showed similarity (between residues 17 and 294) to a 692-amino-acid protein (between amino acids 15 and 258) predicted from a cDNA sequence (Accession No. D63876) (Nagase *et al.*, 1995). This latter protein also showed similarity (at the C-terminus, between amino acids 568 and 689) to the human (and other members) γ-adaptin family of proteins (between residues 710 and 824, Accession No. Y12226). γ-Adaptins are proteins that have been implicated in clathrin-coated vesicle transport, predominantly from the *trans*-Golgi network to the endosome and further to the lysosomes. They also are a part of so called AP1 complex (or adaptor) and are usually composed of 750–850 amino acids (for review see Robinson, 1992; Robinson, 1994; Robinson *et al.*, 1996). Furthermore, another EMBL/GenBank submission (Accession No. AC002400) describes a putative protein that is related to the protein from Accession No. D63876 and displays similar results upon *BlastP*-based comparison with other proteins.

Yet another gene that displays similarity with the N-terminus of *Tom1* was noted. Forty-six ESTs were assembled into a contiguous cDNA sequence (2177 bp, Accession No. AJ010071), capable of encoding a protein of 476 amino acids. This gene represents the human *TOM1* like gene (*TOM1L1*), as its predicted protein sequence displays a high degree of similarity to the *Tom1* gene (37% identity, 55% similarity). The C-terminal portion displayed no similarity to other human proteins deposited in the databases, suggesting that the C-terminus of the *TOM1L1* protein determines its functional specificity.

In summary, the above-described putative domain structure of the TOM1 protein provides the first indication of its normal function. These findings associate the TOM1 protein with proteins implicated in the control of vesicular transport of activated growth factor–receptor complex to lysosomes for degradation, which is a likely mechanism underlying the suppression of cell growth (Komada *et al.*, 1997). The predicted N-terminal vesicle trafficking related domain of TOM1 proteins may be responsible for their endosomal localization. It should be noted that Komada *et al.* (1997) addressed the issue of whether the zinc-finger domain of HGS was responsible for the targeting of this protein to endosomes, and they concluded that it does not perform this function.

Chromosomal Localization of the Mouse Tom1 Gene

Analysis of the genomic sequence of mouse PAC 411O19 indicated that the mouse ortholog of *Hmox1* is adjacent to the *Tom1* gene. Mouse *Hmox1* has previously been mapped to mouse chromosome 10C1 (Saito *et al.*, 1997). However, we failed to confirm this result. Twenty mouse metaphase spreads were analyzed, and specific FISH signals of PAC 411O19 were detected on all of them, including 14 metaphases with double signals on both homologous chromosomes. We determined the localization of the PAC probe by multistep hybridization with mouse painting probes specific for chromosomes 10, 11, 9, 6, and 8. Each step included the identification of two chromosomes with two differentially labeled painting probes (Fig. 4). Chromosomes 6, 9, 10, and 11 were excluded as a localization site of PAC 411O19. The chromosome that carried the PAC signal was painted by chromosome 8-specific probe, and this localization was confirmed by two-color FISH with digoxigenin-labeled PAC 411O19 and Cy3-labeled chromosome 8-specific painting probe. The specific PAC 411O19 signal was detected on both homologs of chromosome 8 in 19 of 20 metaphase spreads. We also determined the localization of the PAC probe at 8C1 by DAPI banding (Fig. 4G). A similar analysis using a chromosome 15-specific painting probe indicated that PAC 411O19 is not a part of the large, neighboring syntenic group of mouse chromosome 15 with human chromosome 22q13–qter (results not shown). These results demonstrate an as yet undescribed syntenic

group between human chromosome 22q13.1 and mouse chromosome 8C1. It is possible that the earlier reported mapping results, which placed the mouse *Hmox1* gene on chromosome 10C1 (Saito *et al.*, 1997), might be due to similar banding patterns of mouse chromosomes 8 and 10 and that *Saito et al.* did not apply in their analysis painting probes to confirm the identity of mouse chromosomes. We recently confirmed the existence of this syntenic group. We characterized an additional human gene located centromeric to *TOM1*, the human *LARGE* (like-acetylglucosaminyl transferase) gene. The mouse ortholog of human *LARGE* is also localized on mouse chromosome 8C1 (Peyrard *et al.*, 1999). Seven syntenic groups between 22q and the mouse genome have previously been reported, involving mouse chromosomes 5, 10, 11, 15, and 16 (Debry and Seldin, 1996). Compared to human chromosome 21, which is of similar size, it is intriguing that human chromosome 22 is divided into many distinct syntenic groups in the mouse genome.

ACKNOWLEDGMENTS

The sequence data for fosmid F77D12, PAC dJ510H16, and BAC bK286B10 were produced by the Human Chromosome 22 Sequencing Group at the Sanger Centre (Hinxton, UK). We thank Dr. Margaret S. Robinson for critically reviewing the manuscript, Dr. Peter G. Zaphiropoulos for the pancreas Marathon-Ready cDNA library, and Dr. Ulf Landegren for advice. This work was supported by grants from the Swedish Cancer Foundation, the Swedish Medical Research Council, the Cancer Society in Stockholm, the Berth von Kantzow Fond, the Ake Wiberg Foundation, Karolinska Hospital, and the Karolinska Institutet to J.P.D. and by grants from NHGRI to B.A.R.

REFERENCES

Asao, H., Sasaki, Y., Arita, T., Tanaka, N., Endo, K., Kasai, H., Takeshita, T., Endo, Y., Fujita, T., and Sugamura, K. (1997). Hrs is associated with STAM, a signal-transducing adaptor molecule. Its suppressive effect on cytokine-induced cell growth. *J. Biol. Chem.* **272:** 32785–32791.

Bean, A. J., Seifert, R., Chen, Y. A., Sacks, R., and Scheller, R. H. (1997). Hrs-2 is an ATPase implicated in calcium-regulated secretion. *Nature* **385:** 826–829.

Burk, O., Worpenberg, S., Haenig, B., and Klempnauer, K. H. (1997). Tom-1, a novel v-myb target gene expressed in amv- and e26-transformed myelomonocytic cells. *EMBO J.* **16:** 1371–1380.

Chen, Y., Hennessy, K. M., Botstein, D., and Tye, B. K. (1992). CDC46/MCM5, a yeast protein whose subcellular localization is cell cycle-regulated, is involved in DNA replication at autonomously replicating sequences. *Proc. Natl. Acad. Sci. USA* **89:** 10459–10463.

Debry, R. W., and Seldin, M. F. (1996). Human/mouse homology relationships. *Genomics* **33:** 337–351.

Deschenes, R. J., Lorenz, L. J., Haun, R. S., Roos, B. A., Collier, K. J., and Dixon, J. E. (1984). Cloning and sequence analysis of a cDNA encoding rat preprocholecystokinin. *Proc. Natl. Acad. Sci. USA* **81:** 726–730.

Dumanski, J. P., Carlbom, E., Collins, V. P., and Nordenskjöld, M. (1987). Deletion mapping of a locus on human chromosome 22 involved in the oncogenesis of meningioma. *Proc. Natl. Acad. Sci. USA* **84:** 9275–9279.

Ellison, J., Dean, M., and Goldman, D. (1993). Efficacy of fluorescence-based PCR-SSCP for detection of point mutations. *BioTechniques* **15:** 684–691.

Fedorova, L., Kost-Alimova, M., Gizatullin, R. Z., Alimov, A., Zabarovska, V. I., Szeles, A., Protopopov, A. I., Vorobieva, N. V., Kashuba, V. I., Klein, G., Zelenin, A. V., Sheer, D., and Zabarovsky, E. R. (1997). Assignment and ordering of twenty-three unique NotI-linking clones containing expressed genes including the guanosine 5′-monophosphate synthetase gene to human chromosome 3. *Eur. J. Hum. Genet.* **5:** 110–116.

Graf, T. (1992). Myb: A transcriptional activator linking proliferation and differentiation in hematopoietic cells. *Curr. Opin. Genet. Dev.* **2:** 249–255.

Hirai, S., and Yaniv, M. (1989). Jun DNA-binding is modulated by mutations between the leucines or by direct interaction of fos with the TGACTCA sequence. *New Biol.* **1:** 181–191.

Komada, M., Masaki, R., Yamamoto, A., and Kitamura, N. (1997). Hrs, a tyrosine kinase substrate with a conserved double zinc finger domain, is localized to the cytoplasmic surface of early endosomes. *J. Biol. Chem.* **272:** 20538–20544.

Kutty, R. K., Kutty, G., Rodriguez, I. R., Chader, G. J., and Wiggert, B. (1994). Chromosomal localization of the human heme oxygenase genes: Heme oxygenase-1 (HMOX1) maps to chromosome 22q12 and heme oxygenase-2 (HMOX2) maps to chromosome 16p13.3. *Genomics* **20:** 513–516.

Lekanne Deprez, R. H., Riegman, P. H., van Drunen, E., Warringa, U. L., Groen, N. A., Stefanko, S. Z., Koper, J. W., Avezaat, C. J., Mulder, P. G., Zwarthoff, E. C., and Hagemeijer, A. (1995). Cytogenetic, molecular genetic and pathological analyses in 126 meningiomas. *Eur. J. Hum. Genet.* **54:** 224–235.

Lohi, O., Poussu, A., Merilainen, J., Kellokumpu, S., Wasenius, V. M., and Lehto, V. P. (1998). East, an epidermal growth factor receptor- and eps15-associated protein with src homology 3 and tyrosine-based activation motif domains. *J. Biol. Chem.* **273:** 21408–21415.

Lupas, A., Van, D. M., and Stock, J. (1991). Predicting coiled coils from protein sequences. *Science* **252:** 1162–1164.

Monla, Y. T., Peleg, S., and Gagel, R. F. (1995). Cell type-specific regulation of transcription by cyclic adenosine 3′,5′-monophosphate-responsive elements within the calcitonin promoter. *Mol. Endocrinol.* **9:** 784–793.

Nagase, T., Seki, N., Tanaka, A., Ishikawa, K., and Nomura, N. (1995). Prediction of the coding sequences of unidentified human genes. IV. The coding sequences of 40 new genes (KIAA0121–KIAA0160) deduced by analysis of cDNA clones from human cell line KG-1. *DNA Res.* **2:** 167–174.

Orita, M., Iwahana, H., Kanazawa, H., Hayashi, K., and Sekiya, T. (1989). Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. *Proc. Natl. Acad. Sci. USA* **86:** 2766–2770.

Orita, M., Suzuki, Y., Sekiya, T., and Hayashi, K. (1989). Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. *Genomics* **5:** 874–879.

Paul, R., Hu, B., Musahl, C., Hameister, H., and Knippers, R. (1996). Coding sequence and chromosome mapping of the human gene (CDC46) for replication protein hCdc46/Mcm5. *Cytogenet. Cell Genet.* **73:** 317–321.

Peyrard, M., Seroussi, E., Sandberg-Nordqvist, A. C., Fransson, I., Xie, Y.-G., Han, F.-Y., Collins, J., Dunham, I., Kost-Alimova, M., Imreh, S., and Dumanski, J. P. (1999). The human *LARGE* gene (***L***ike-***A***cetyl***G***lucosaminylt***R***ansferas***E***) from 22q12.3–q13.1 is a new, distinct member of the glycosyltransferase gene family. *Proc. Natl. Acad. Sci. USA* **96:** 598–603.

Robinson, M. S. (1992). Adaptins. *Trends Cell Biol.* **2:** 293–297.

Robinson, M. S. (1994). The role of clathrin, adaptors and dynamin in endocytosis. *Curr. Opin. Cell Biol.* **6:** 538–544.

Robinson, M. S., Watts, C., and Zerial, M. (1996). Membrane dynamics in endocytosis. *Cell* **84:** 13–21.

Ruttledge, M. H., Sarrazin, J., Rangaratnam, S., Phelan, C. M., Twist, E., Nordenskjöld, M., Collins, V. P., Dumanski, J. P., and Rouleau, G. (1994). Evidence for the complete inactivation of the neurofibromatosis type 2 gene in the majority of sporadic meningiomas. *Nat. Genet.* **6:** 180–184.

Ruttledge, M. H., Xie, Y.-G., Han, F.-Y., Peyrard, M., Collins, V. P., Nordenskjöld, M., and Dumanski, J. P. (1994). Deletions on chromosome 22 in sporadic meningioma. *Genes Chromosomes Cancer* **10:** 122–130.

Saito, O. F., Ikeuchi, T., Matsumoto, M., and Kurata, S. (1997). Assignment of the mouse heme oxygenase genes: Heme oxygenase-1 (Hmox1) to chromosome 10 band C1 and heme oxygenase-2 (Hmox2) to chromosome 16 band B1. *Cytogenet. Cell Genet.* **77:** 180–181.

Sandberg-Nordqvist, A.-C., von Holst, H., Holmin, S., Sara, V. R., Bellander, B. M., and Schalling, M. (1996). Increase of insulin-like growth factor (IGF)-1, IGF binding protein-2 and -4 mRNAs following cerebral contusion. *Brain Res. Mol. Brain Res.* **38:** 285–293.

Seizinger, B. R., de la Monte, S., Atkins, L., Gusella, J. F., and Martuza, R. L. (1987). Molecular genetic approach to human meningioma: Loss of genes on chromosome 22. *Proc. Natl. Acad. Sci. USA* **84:** 54l9–5423.

Wolf, E., Kim, P. S., and Berger, B. (1997). MultiCoil: A program for predicting two- and three-stranded coiled coils. *Protein Sci.* **6:** 1179–1189.

Zang, K. D., and Singer, H. (1967). Chromosomal constitution of meningiomas. *Nature* **216:** 84–85.

Zipfel, P. F., Decker, E. L., Holst, C., and Skerka, C. (1997). The human zinc finger protein EGR-4 acts as autoregulatory transcriptional repressor. *Biochim. Biophys. Acta* **1354:** 134–144.