Mammalian Genome 12, 852–861 (2001).
DOI: 10.1007/s00335-001-3037-3

Mammalian Genome
Incorporating Mouse Genome

MACF1 gene structure: a hybrid of plectin and dystrophin

Tzy-Wen L. Gong, Cagri G. Besirli, Margaret I. Lomax

Kresge Hearing Research Institute, Department of Otolaryngology/Head Neck Surgery, 9301E MSRB III, 1150 W. Medical Center Dr., Box 0648, University of Michigan, Ann Arbor, MI 48109-0648, USA

Received: 24 April 2001 / Accepted: 29 June 2001

Abstract. Mammalian MACF1 (Macrophin1; previously named ACF7) is a giant cytoskeletal linker protein with three known isoforms that arise by alternative splicing. We isolated a 19.1-kb cDNA encoding a fourth isoform (MACF1-4) with a unique N-terminus. Instead of an N-terminal actin-binding domain found in the other three isoforms, MACF1-4 has eight plectin repeats. The MACF1 gene is located on human Chr 1p32, contains at least 102 exons, spans over 270 kb, and gives rise to four major isoforms with different N-termini. The genomic organization of the actin-binding domain is highly conserved in mammalian genes for both plectin and BPAG1. All eight plectin repeats are encoded by one large exon; this feature is similar to the genomic structure of plectin. The intron positions within spectrin repeats in MACF1 are very similar to those in the dystrophin gene. This demonstrates that MACF1 has characteristic features of genes for two classes of cytoskeletal proteins, i.e., plectin and dystrophin.

The plakins, a family of large structural proteins that organize cytoskeletal elements, include plectin, dystonin/bullous pemphigoid antigen 1 (BPAG1), desmoplakin, envoplakin (Ruhrberg and Watt 1997), periplakin (Ruhrberg et al. 1997), and macrophin1 (MACF1), which was previously named ACF7 for actin cross-linking protein family 7 (Bernier et al. 1996). These proteins have similar modular structures composed of globular N-terminal and C-terminal domains that are connected by a central, α-helical, coiled-coil, rod-like domain. Both plectin and dystonin/BPAG1n (the neural isoform of BPAG1) contain an N-terminal actin-binding domain (ABD), whereas desmoplakin, envoplakin, and periplakin do not. Since these proteins provide structural support in epidermal cells and neurons, loss of plakin function often results in diseases of the skin as well as neural tissues (Fuchs et al. 1997; Mahoney et al. 1998). For example, mutations in the human plectin gene cause epidermolysis bullosa simplex with muscular dystrophy (EBS-MD; McLean et al. 1996). Mutation of the mouse gene for dystonin/BPAG1 weakens mechanical integrity of the stratified epithelium, resulting in dystonia musculorum with severe neuronal degeneration in mice (Brown et al. 1995; Dalpe et al. 1999; Dowling et al. 1997; Guo et al. 1995).

MACF1 is classified as a member of the plakin family because its N-terminal ABD is more similar to the ABD in plectin and dystonin/BPAG1n than that in α-actinin/β-spectrin/dystrophin (Bernier et al. 1996; Byers et al. 1995). MACF1 was first identified as a partial human cDNA and named ACF7 by Kunkel and colleagues in a search for new types of actin-binding proteins related to dystrophin (Byers et al. 1995). In a similar screen for actin-binding proteins related to dystonin, Kothary and colleagues identified three isoforms of mouse ACF7 with different N-termini (Bernier et al. 1996). Full-length cDNAs of one of these isoforms have been cloned subsequently in human and mouse and given different names, e.g., macrophin (Okuda et al. 1999), trabeculin-α (Sun et al. 1999), and MACF (Microtubule Actin Cross-linking Factor; Leung et al. 1999). As with plectin and BPAG1n (Brown et al. 1995; Fuchs et al. 1999), alternative splicing and promoter usage appear to result in the multiple MACF1 isoforms. MACF1 isoforms 1 and 2 (MACF1-1 and MACF1-2; Bernier et al. 1996; Leung et al. 1999) have identical ABDs, but different 5′ UTRs and N-terminal sequences preceding the ABD. Isoform 3 (MACF1-3), in contrast, contains a unique 5′UTR, a longer N-terminal sequence, and just the second half of the ABD (Bernier et al. 1996). The sequences following the ABD appear to be otherwise identical, but have not been characterized completely.

In this report we present the molecular cloning of MACF1-4, a fourth isoform of MACF1. The 19.1-kb full-length human MACF1-4 cDNA did not encode an ABD, but instead contained plectin repeats at the N-terminus. The N-terminus of this isoform is clearly distinct from any of the three mammalian MACF1 isoforms cloned previously. However, the remaining 13 kb of the 19.1-kb cDNA was virtually identical to the previously characterized full-length human MACF1 cDNAs (Okuda et al. 1999; Sun et al. 1999) and was highly homologous to the mouse ortholog Macf1 (Leung et al. 1999). Analysis of a BAC/PAC contig covering the MACF1 gene identified exons encoding all four isoforms and provided insight into mechanisms that direct production of alternative N-terminal sequences of MACF1. The MACF1 gene appears to be a hybrid derived from ancestral genes for plectin and dystrophin.

Materials and methods

Materials. Routine restriction and modifying enzymes were obtained from commercial sources and used according to the manufacturer's instructions. SuperScript II reverse transcriptase was from Life Technologies (Gaithersburg, Md.). Advantage cDNA Amplification System, the human RNA master dot blot, and the human cDNA library screening panel were from Clontech (La Jolla, Calif.). Human pituitary and HepG2 cDNA libraries in λZap were from Stratagene (Catalog No. 936309; La Jolla, Calif.), and a human heart cDNA library was from OriGene (Rockville, Md.). The pGEM-T EASY Cloning and In Vitro Transcription Systems were from Promega (Madison, Wis.). Strip-EZ RNA probe synthesis and removal kit was from Ambion (Austin, Tex.). Human and mouse IMAGE clones, human CITB BAC DNA Pools (combined B and C libraries derived from 987SK cells and sperm; release IV, Cat. No. 96011) and individual clones in the pADSacBII vector were from Research Genetics (Huntsville, Ala.). Two P1 artificial chromosome (PAC) clones in the pADSacBII vector derived from male foreskin fibroblasts were from Genome Systems (St. Louis, Mo.).

Molecular cloning of human MACF1-4 cDNA. A chick partial cDNA, KH124, was isolated in a differential display experiment designed to identify genes differentially expressed during regeneration of the auditory epithelium after noise trauma (Gong et al. 1996). A BLAST (Altschul et al. 1997) search of dbEST with KH124 identified several overlapping

MACF1 for macrophin1 is an approved symbol by the Human and Mouse Nomenclature Committees. The nucleotide sequence data reported in this paper have been submitted to GenBank and have been assigned the accession numbers AF317696 (human MACF1-4 cDNA) and AF325326–AF325343 (genomic sequences of human MACF1).

Correspondence to: M.I. Lomax; E-mail: mlomax@umich.edu

human and mouse ESTs and subsequently KIAA0465 (GenBank AB007934). We obtained four human and mouse cDNA clones for further analysis: two human IMAGE clones KH327 (IMAGE CloneID 198389) and KH362 (CloneID 594081) overlapped and generated a 3-kb contig corresponding to the most 3′ end of cDNA. Five additional MACF1 cDNAs were obtained by screening 10^6 phage clones from a human pituitary gland cDNA library (Stratagene), initially with the 5′ end of KH327 (KH408 and KH418), then with the 5′ end of each longer human cDNA (KH442, KH454, and KH478). These additional cDNAs extended the MACF1 contig to 13.4-kb. RT-PCR was performed on total RNA of human EBV-transformed lymphoblastoid cells to show that the contig up to this point was indeed derived from the same transcript. As no additional cDNAs could be isolated from this pituitary gland library, we screened a human heart cDNA library with the 5′ end of KH478, which identified KH515 and extended the contig by an additional 2.8-kb to a total of 16.2 kb. The remaining 5′ end of the MACF1-4 cDNA was cloned from a genomic BAC clone (CITB-372N10) with a continuous ORF.

Isolation of genomic DNA for human MACF1 gene. Human MACF1 genomic clones were isolated, starting at the 3′ end of the gene (Fig. 2A). PCR assays were used to identify genomic PAC and BAC clones for human MACF1 gene. Gene-specific primers were designed to the MACF1 cDNA sequences. Two PAC clones (clones 144E1 and 180I11) corresponding to the 3′ end of human MACF1 gene were identified by Van Camp's group as part of a contig for the DFNA2 candidate gene region. By using primers from the 5′ ends of MACF1 cDNAs, we identified at least 14 MACF1 BAC clones and obtained and sequenced 9 of these: 157P10 (KH526), 336H16 (KH527), 372N10 (KH528), 519O8 (KH530), 210A12 (KH674), 367P6 (KH675), 409D1 (KH676), 461A22 (KH677), and 540L7 (KH678). DNA was prepared from 0.5–1 L saturated liquid cultures with KB-100 Magnum Purification kit (Genome Systems) or Qiagen Plasmid Maxi Kit with an additional treatment with RNaseA (10 μg/ml) to remove trace amounts of RNA. Purified BAC/PAC DNAs were subjected to sequencing and PCR analysis.

DNA and protein sequence analysis. Purified plasmid and BAC/PAC DNAs were sequenced, initially with vector primers, then with gene-specific primers, at the University of Michigan DNA Sequencing Core on an automated DNA sequencer by using ABI Taq DyeDeoxy terminator cycle sequencing. A few intronic regions that were difficult to sequence directly from BAC/PAC clones were PCR amplified, subcloned, and sequenced. If sequencing difficulty persisted, size estimates of these regions were obtained, based on PCR. DNA sequences were aligned to form a cDNA contig and 19 genomic contigs with the SEQMAN program of LaserGene (DNASTAR, Madison, Wis.). Database searches with either DNA or protein sequences were performed with the BLAST search program (Altschul et al. 1997) on NCBI databases (http://ncbi.nlm.nih.gov). Protein structure analysis for domain matches, e.g., plectin repeat, spectrin repeats, GAR/GAS domain, and EF hands, was performed with Pfam (release 5.4) (http://pfam.wustl.edu/hmmsearch.shtml).

SSCP analysis. One pair of primers (8622F and 6441I) corresponding to the 3′UTR of the mouse Macf1 cDNA was used to amplify by PCR genomic DNA from mouse strains C57BL/6J and *Mus spretus.* Nucleotide sequencing identified a single-base-pair mismatch as well as size difference in amplified products between these two strains (261 bp in C57BL/6J vs 262 bp in *M. spretus*). SSCP analysis of the amplified region was then performed on the parental strains and 94 progeny from the BSS interspecific backcross (C57BL/6J × *M. spretus*) F_1 × *M. spretus* from The Jackson Laboratory (Rowe et al., 1994). SSCP analysis was performed as described previously (both primers were end labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase. PCR was performed in a 10-μl reaction containing 50 ng genomic DNA, 100 μM dNTP, 300 nM unlabeled primers, 100 nM end-labeled primers, 0.5 unit Taq polymerase (Sigma), 10 mM Tris (pH 8.0–8.3), 5 mM NH_4Cl, 25 mM KCl, and 1.5 mM $MgCl_2$. Thermocycling conditions were: 95°C for 15 s, 60°C for 30 s, followed by 72°C for 1 min, for a total of 35 cycles. PCR reactions were denatured at 95°C for 10 min, cooled immediately on ice, and loaded on 0.4 × MDE gels (Mutation Detection Enhancement gel, FMC BioProducts) by using 0.7 × TBE buffer for electrophoresis at 12 W for 18 h. Gels were dried and exposed to Kodak X-Omat film overnight at room temperature. C57BL/6J-specific SSCP bands were scored for haplotype analysis. The data are deposited and available on The Jackson Laboratory web page (http://www.jax.org/resources/documents/cmdata).

RNA blot hybridization. Isolation of total RNA from 4-week-old mice (strain C57BL/6J), electrophoretic separation, transfer to Nytran membranes, preparation of an anti-sense riboprobe, and the hybridization conditions were as described previously (Gong et al. 1999). Agarose (1%)-formaldehyde gels were run for 8 h at 57 V. The mouse tissue Northern blot was hybridized with an [α-^{32}P]UTP-labeled anti-sense riboprobe derived from KH331 (1339-bp insert corresponding to the 3′-UTR of the mouse Macf1 cDNA) at 65°C for 22 h. The most stringent post-hybridization wash was in 0.2× SSC–0.5% SDS at 65°C. RNA sizes were based on extrapolated estimates from RNA molecular weight markers (size ranges of 1.5 kb–6.9 kb) and a 10-kb Kif1bp204 transcript studied previously (Gong et al. 1999). The KH331 anti-sense riboprobe was labeled with [α-^{32}P]UTP and modified CTP (StripEZ RNA synthesis kit) for analysis of a master dot blot (Clontech) containing poly(A)+ RNA from 50 different human tissues. Hybridization conditions were similar to those used for the mouse tissue Northern blot, but the most stringent wash was in 0.5× SSC–0.5% SDS at 60°C to allow for cross-species hybridization. Membranes were exposed to Kodak X-Omat films with intensifying screens at −80°C for 2 days (human dot-blot) or 3 days (mouse Northern).

Results

Molecular cloning of MACF1-4. We have identified and cloned a cDNA encoding a new isoform of MACF1 that lacks an actin-binding domain (ABD). The cDNA sequence was generated by screening human pituitary and heart cDNA libraries (Fig. 1A), starting with a probe from the 3′untranslated region (UTR), as well as by sequencing the large N-terminal exon from a BAC clone. The composite 19.1-kb cDNA sequence contained a 0.1-kb 5′UTR, 17.8-kb ORF, and a 1.2-kb 3′ UTR (Fig. 1A), and predicted a protein of 670 kDa. This isoform of MACF1 is one of the largest proteins identified to date. Pfam analysis predicted eight plectin repeats in the N-terminus and 27 spectrin repeats in the central region, followed by two putative Ca-binding EF hands, a distinctive Gas2 homology domain, and a serine/glycine-rich C-terminus. The 3′-end 13 kb of the MACF1-4 sequence, encoding 26 of 27 spectrin repeats and the C-terminal regions, was virtually identical to full-length sequences of human MACF1-2, macrophin (Okuda et al. 1999), and trabeculin-α (Sun et al. 1999), and was highly similar to its mouse ortholog, Macf1 (Leung et al. 1999). The first 5 kb of the cDNA, however, was unique and shared no significant homology with any mammalian MACF1 cloned previously (Bernier et al. 1996; Byers et al. 1995; Leung et al. 1999; Okuda et al. 1999; Sun et al. 1999; Fig. 1B). The MACF1-4 cDNA does not contain any sequence for the ABD present in the other three isoforms despite many attempts by RT-PCR to link its 5′-end sequence to the ABD sequence. Instead, MACF1-4 has a unique N-terminus containing 8 plectin repeats and a spectrin repeat not present in the other isoforms. The plectin repeat, defined here based on the Pfam algorithm, is a 45-amino acid long α-helical domain; it was first identified in plectin (Wiche et al. 1991) and is also present in other plakins. The spectrin repeat is an approximately 109 amino acid helix-turn-helix repeat present in spectrin, dystrophin, and other members of the spectrin subfamily of actin crosslinking proteins (Winder et al. 1995).

Mapping the mouse Macf1 gene. To determine the chromosomal location of the orthologous mouse gene, we used an SSCP assay for the 3′ UTR of the corresponding mouse cDNA on the Jackson mouse BSS backcross (Rowe et al. 1994). Linkage and haplotype analyses revealed that this gene is located on mouse Chr 4 at position 55.6 (Mouse Chromosome Committee) with the following gene order: *Dab1/Iapls3-10/Pmv19*–(1.06 ± 1.06 cM)–*D4Mit120*–(1.06 ± 1.06 cM)–*Faah/Mf2/Mp1*–(1.06 ± 1.06 cM)–*D4Mit334*–

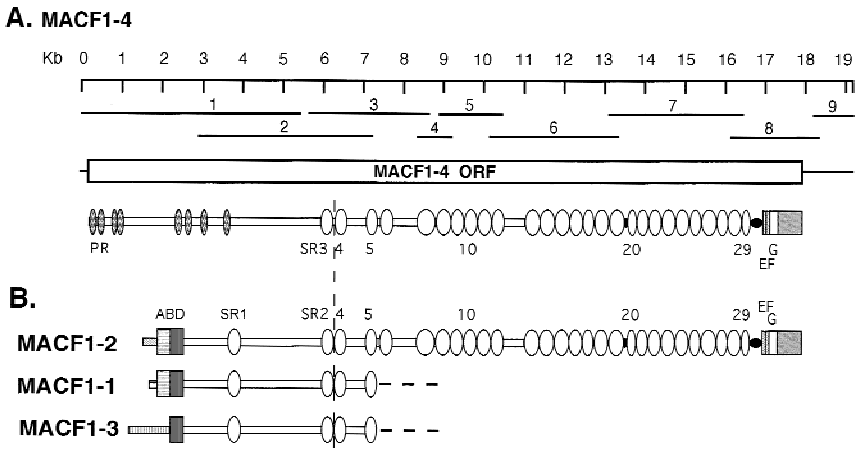


Fig. 1. Human cDNA and predicted domain structure for MACF1 isoform 4 (MACF1-4). **(A)** The MACF1-4 cDNA contig, ORF, and graphic representation of the predicted protein domains. The domains preceding the vertical dotted line are unique to MACF1-4; the domain structure following this dotted line is common to all four isoforms. Abbreviations and symbols used: ABD, actin-binding domain (darker shaded boxes); PR, plectin repeat (dotted ovals); SP, spectrin repeat (open ovals); EF, EF hand (rectangles with backward diagonal stripes); G, Growth-Arrest-Specific Protein 2 (Gas2) domain (open rectangle). Numbered horizontal lines indicate the position and extent of DNA sequences used to construct the MACF1-4 cDNA sequence (GenBank No. AF317696). Sequence 1 was derived from a continuous ORF in exon 38. Sequence 2, a partial cDNA isolated from a human heart cDNA library (OriGene), extended from nt 2951 (in exon 38) to nt 7100, the sequence following spectrin repeat 5 of the common region. Sequences 3–7 were partial cDNAs isolated from a human pituitary cDNA library. Sequences 8 and 9 were partial cDNAs from IMAGE collections (CloneID 198389 and 594081, respectively). These two clones were identified through database searches with a chick cDNA KH124 from a previous study (Gong et al. 1996). The open box represents the MACF1-4 coding region; the lines preceding and following the open box represent the 5′UTR and 3′UTR, respectively. Pfam analysis of MACF1-4 protein sequence identified 8 plectin repeats, 27 spectrin repeats (SR 3-29), 2 putative Ca-binding EF hands, a Gas2 domain (Collavin et al. 1998), and a glycine/serine-rich C-terminus (diagonal stripes). **(B)** Structural domains of MACF1 isoforms 1–3. Each isoform (MACF1-1, -2, and -3) contains an isoform-specific N-terminus (shaded bar), followed by either a complete actin-binding domain in isoforms 1 and 2 or only the second half of the ABD in isoform 3. Pfam analysis predicted 28 spectrin repeats (SR1, 2, and 4-29) in the full-length MACF1-2 (macrophin, GenBank AB029290), the only other MACF1 isoform for which a complete cDNA contig has been generated. The N-terminal structures of MACF1-1 and MACF1-3 are based on BLAST alignments of the mouse partial cDNA sequences (U67203 and U67205, respectively; Bernier et al. 1996) with human genomic sequences (see Figs. 2 and 3).

(1.06 ± 1.06 cM)–*Iapls3-31/Iapls1-19*–(1.06 ± 1.06 cM)–*Macf1*/*Aclp7*/*Bmp8a*/*D4Mit11*–(1.06 ± 1.06 cM)–*D4Mit12*–(1.06 ± 1.06 cM)–*Prp18*–(1.06 ± 1.06 cM)–*D4Mit16*–(1.06 ± 1.06 cM)–*Matn1*/*Pmp22-rs/Ptpro*. This region of mouse Chr 4 shares a region of conserved synteny with human Chr 1p32, where the human MACF1 has been mapped (WI-13856, STS corresponding to 3′ UTR of MACF1; GenBank G24302). These mapping data thus confirmed the locations of the orthologous genes: human MACF1 and mouse *Macf1*. Kothary and colleagues previously mapped the 5′ end of the mouse *Macf1* gene (locus designation *Aclp7*) by RFLP analysis to the same location on mouse Chr 4, by using the same Jackson BSS backcross panel (Bernier et al. 1996). These results showed a close link between the 3′ end of our new cDNA and the 5′ end of the mouse *Acf7* cDNA. The genomic organization of the human MACF1 gene (see below) confirmed that MACF1 and ACF7 were indeed the same gene.

Genomic organization of human MACF1. To determine the exon-intron structure of MACF1, overlapping human BAC and PAC clones spanning the majority of the MACF1 gene were isolated by PCR screening and sequenced as described in the Methods. Genomic clones for the first two exons were identified from unordered, partially finished sequences in the high throughput genome sequence (htgs) database. We identified 102 exons in the MACF1 gene by comparing cDNA and genomic sequences (Fig. 2A). All exons were flanked by consensus acceptor and donor sequences. The exon sizes range from 9 bp (exon 72) to over 5.4 kb (exon 38). The 5′-most BAC (RP4-648J17; Fig. 2A) was identified in htgs sequences and provided the sequence surrounding exons 1 and 2. This BAC clone also contains the NDUFS5 gene encoding the 15-kDa subunit of mitochondrial NADH-coenzyme Q reductase and therefore defines a 5′ telomeric neighbor of the MACF1 gene. The 3′-most PAC (RPCI-1 144E1) contains several STS markers derived from the 3′UTR of MACF1. Another htgs BAC clone RP11-161O10 (GenBank AC024511, containing intron 44 to exon 102 of MACF1) overlaps with BAC clone RP11-118J21 (GenBank AL033527), which contains BMP8 (bone morphogenic protein 8) and MYCL1 (L-MYC) genes. This BAC/PAC contig therefore contains the entire human MACF1 gene, which is located on Chr1p32 between NDUFS5 (telomeric) and BMP8/MYCL1 (centromeric) genes.

The overall order of exons encoding the isoform-specific regions of MACF1-1, 2, and 3 and the ABD is quite similar to the organization of the murine plectin gene (*Plec1*) (Fuchs et al. 1999; Liu et al. 1996). MACF1 exons 1 and 2 encode the unique 5′UTRs and isoform-specific N-termini of isoforms 2 and 1, respectively. These alternative exons are then spliced to exon 3, the exon encoding the first part of the ABD (Fig. 2B). The MACF1 ABD is encoded by seven exons: exons 3–6 for the first half, and exons 8–10 for the second half of the ABD. Exons 3–6 are used in MACF1-1 and -2, but not MACF1-3. Exon 7 contains the MACF1-3 specific sequence, which is spliced into exon 8 to generate the incomplete ABD characteristic of isoform 3. Exons 8–10 encode the second half of the ABD, which is present in all three isoforms (Fig. 2B).

Exons 1, 2, and 7 encode the 5′UTRs and initiation Met codons of isoforms 2, 1, and 3, respectively. These sequences enabled us to define more precisely the N-termini of MACF1 isoforms, which are presented with the sequence of the predicted 5′ UTRs in Fig. 3A–C. We identified a putative initiation ATG codon for MACF1-2 in exon 1 (Fig. 3A). This codon and an in-frame stop codon 145 bp upstream, with no other intervening ATG codons in-frame, are preceded by a good match (GCCtggGCCATG) to the Kozak consensus sequence (GCCRCCATG) (Kozak 1987). The predicted N-terminal amino acid sequence of exon 1 is identical to that of human macrophin and mouse Macf sequences. The second half of the sequence is also very similar to exon A (77% identity) encoding isoform 2-specific sequences of human and mouse genes for dystonin/Bpag1n (Bernier et al. 1996), as well as exon 1c of the plectin gene (57% identity) (Fuchs et al. 1999). Although human MACF1-1 cDNA had not been cloned, exon prediction was possible based upon mouse Macf1-1 partial cDNA sequence. Exon 2 of the human MACF1 gene contained a potential ATG codon (Fig. 3B). This is the first in-frame start codon following a stop codon

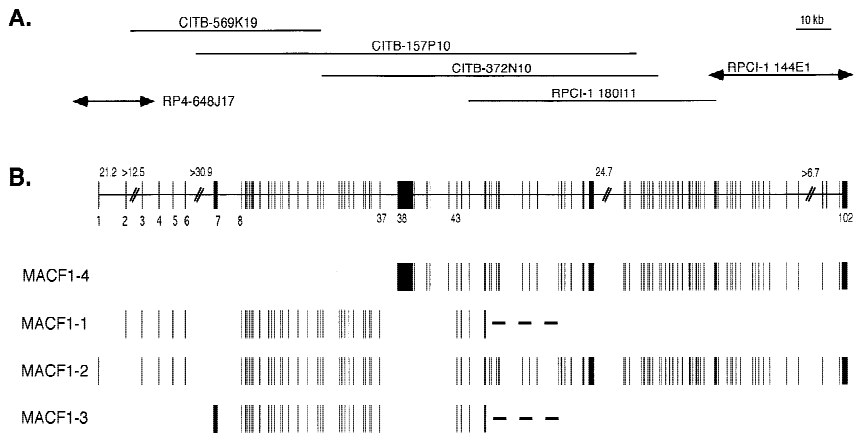


Fig. 2. Genomic organization of the human MACF1 gene. **(A)** BAC/PAC contig covering the MACF1 gene. Horizontal lines indicate the name, position and extent of representative BAC and PAC genomic clones spanning the MACF1 gene on human Chr 1. Clone overlaps were confirmed by marker content analysis. Arrows indicate that the positions of the BAC/PAC ends have not been determined. Twelve BAC clones were isolated from a CITB genomic BAC library by PCR assays, using primers derived from MACF1-4 and MACF1-2 cDNA sequences. Two PAC clones were from an RPCI-1 library identified by PCR screening against PAC contigs for the DFNA2 candidate gene region by Dr. Guy Van Camp. Clone RP4-648J17 (GenBank AL139015), identified by BLAST searches with mouse MACF1-1 (U67203) and human MACF1-2 cDNA sequences (AB029290), contained exons 1 and 2. This clone also contained the NDUFS5 gene (NADH dehydrogenase Fe-S protein 5, also known as NADH ubiquinone oxidoreductase 15 kDa IP subunit) and defines the telomeric neighbor of MACF1. Clone RP11-161O10 (GenBank AC024511) extends from intron 44 to exon 102 of MACF1 and overlaps with clone RP11-118J21 (AL033527), which contains BMP8 (bone morphogenetic protein 8) and MYCL1 and defines the centromeric neighbors of the MACF1 gene. **(B)** Exon-intron organization of the human MACF1 gene. Horizontal lines denote introns (minimum sizes of large introns noted above lines). Vertical lines and black boxes represent exons (numbered below), starting with exon 1, which contains the 5′UTR and start codon of MACF1-2. Exons contained in the MACF1-4 cDNA sequence are represented graphically below the gene, as are exons represented in the published cDNA sequences of the other three MACF1 isoforms, i.e., full-length human MACF1-2 cDNA (AB029290) and partial mouse cDNAs for MACF1-1 (U67203) and MACF1-3 (U67205). Dotted lines in MACF1-1 and MACF1-3 indicate that these regions have not been cloned and are assumed to be identical to isoform 2.

(tga) 531 nts upstream. The sequence preceding this Met codon (GCgGCCATG) is also a good match to the Kozak consensus sequence. The predicted MACF1-1 N-terminus is 15 aa longer than the published mouse Macf1-1 partial sequence. The human MACF1-1 N-terminus is also similar to human and mouse dystonin isoform 1 encoded by exon A′, as well as to plectin encoded by exon 1f (81% and 62% identity, respectively, in the region of available sequence). Although the sequence preceding our predicted start codon in exon 7 for MACF1-3 (aagGaaATG) does not match the Kozak consensus, the mouse Macf1-3 cDNA predicts the same initiation Met codon (Fig 3C). In addition, no other ATG codons are present in exon 7 or immediate downstream. The human and mouse sequences of the isoform 3-specific regions shared only 76% amino acid identity, in contrast to the high sequence identity between human and mouse MACF1 isoforms 1 as well as 2. Additional sequences following the ABD and common to isoforms 1, 2, and 3 are encoded by exons 11–37.

Exon 38, the largest exon of the MACF1 gene, encodes the MACF1-4 isoform-specific N-terminus (Fig. 3D). This exon contained an ATG codon preceded by (aaaGCCATG), a reasonable match to the Kozak consensus sequence. Although there is a potential 3′ splice acceptor sequence upstream, we found no evidence for any mRNA in which upstream exons splice into exon 38. This analysis included repeated attempts to link exon 38 to either exon 37 or one of the exons for the ABD either by PCR of cDNA libraries or by RT-PCR of human mRNA derived from several tissues, including heart, lung, and brain. We therefore propose this ATG as the initiation codon for isoform 4. Pfam analysis identified eight plectin repeats that are all encoded by exon 38 and are present only in MACF1-4 (Fig. 1B; 4B). The first plectin repeat is 12-amino acids downstream of this ATG. Additional isoform 4-specific sequences were encoded by exons 39–42, which were not present in isoforms 1, 2, or 3.

Exons 19–37 and 41–91 encode spectrin repeats that contribute to the central rod domain in MACF1. Pfam analysis identified 29 spectrin repeats (SR) with E values < 0.05 (Fig. 4A). Spectrin repeats 1 (SR1; exons 19–20) and 2 (SR2; exons 35–37) are present only in isoforms 1, 2, and 3. SR3 (exons 41–42) is present only in MACF1-4. Exons 43–91 encode 26 additional spectrin repeats, which are present in both full-length MACF1-2 and MACF1-4. Thus, MACF1-2 contains 28 repeats (SR 1,2, and 4–29), whereas MACF1-4 contains 27 repeats (SR 3–29). Spectrin repeats were generally encoded by small exons interrupted by small introns. We aligned all 29 repeats and noted the positions of introns within the repeats (Fig. 4B). The most striking feature of this analysis was the nearly complete conservation of an intron position in the middle of helix B of the repeat. Only exon 59, the largest SR-encoding exon (1472 bp) in human MACF1, lacks this conserved intron and codes for two repeats (SR 10 and 11). This exon-intron organization, including the lack of the intron in two spectrin repeats, is strikingly similar to that of the spectrin/dystrophin repeats identified in human dystrophin (Koenig and Kunkel 1990; Winder et al. 1995).

The remaining exons (95–102) give rise to protein sequences that distinguish MACF1 from other plakins. Exons 95 and 96 encoded two putative EF hands, exon 99 the Gas2 domain, and exons 100–102 the serine/glycine-rich C-terminus. Exon 102, the last exon (>1.4 kb), also contained the translation stop codon TAA and a 3′ UTR.

Alternative splicing occurs not only at the 5′ end to generate four major isoforms, but also throughout the gene. Exon 54 (63 bp) is present in our cDNA sequence, but not in human trabeculin-α (MACF1-2). Exon 98 (18 bp), immediately preceding the predicted Gas2 domain, is not present in macrophin MACF1-2 or another sequence, KIAA0465. Furthermore, exon 100 (111 bp) was identified as a partial sequence of three ESTs derived from breast and colon cDNA libraries, but was not present in any other full-length cDNA cloned. As the lengths of these exons are multiples of 3, alternative splicing would not cause a frame-shift of translation. The functional significance of such alternative splicing is not clear. In mouse Macf1-2 cDNA, there are no sequences corresponding to exons 64 (327 bp) or 72 (9 bp) in the human gene. It is not clear whether this is the result of alternative splicing or missing exons. Another discrepancy between the human and mouse genes is a 12-bp sequence present in the mouse cDNA, but not in human exon 94, nor as a small exon in the human gene.

Expression of MACF1 in human and mouse tissues. To examine the overall expression pattern of MACF1 in human tissues, we hybridized a dot blot containing poly(A)+ RNA from 50 different human tissues, including sub-regions of the brain, with a mouse 3′

A. EXON 1 ORF prediction

```
tcccattttcaggtgaagaaactgaggtacagataagcaggactagaaccctctgcttctg
Cttcccaagtcccagtgttccttatgggaccccttttgtctgctcacattcacccctttc
tccttcagatcacttctccctgggctcccaggcctcctgcagcagccccgcctgggcc
ATGTCTTCCTCAGATGAAGACGCTCAGTGAGCGGTCATGTCGGAGTGAGCGGTCTTGT
 M  S  S  S  D  E  E  T  L  S  E  R  S  C  R  S  E  R  S  C
CGGAGTGAGCGATCTTACAGGAGCGAGCGGTCGGGGAGCCTGTCTCCCTGTCCCCCAGGG
 R  S  E  R  S  Y  R  S  E  R  S  G  S  L  S  P  C  P  P  G
GACACCTTGCCCTGGAACCTGCCACTGCATGAGCAGAAAAAGCGGAAAAGCCAGGATTCG
 D  T  L  P  W  N  L  P  L  H  E  Q  K  K  R  K  S  Q  D  S
GTGCTGGACCCTGCAGAGCGTGCTGTGGTCAGAGTCGCTGgtaag
 V  L  D  P  A  E  R  A  V  V  R  V  A
```

B. Exon 2 ORF prediction

```
ctttcatcctgacgcgcggggcctcccacccagcgggactcccgcgtgggccggcctctc
Gcgcccctcggctcgggccccagtccgggccgggcgggggtcggcagccccctgggggacc
Cgtgtgggcagcccgggcccagccggccgcagcctggggccgaggactcgccgggacccg
Gaggcggcggggagaggggcggggaacgggccgggcctggcgctcctgacaggaggagc
Cgccgccgccgcccgccgctgcagccgcgccgggggcgggctgagggaggagcggagccga
Ggggtgaggacgcggaaacgcgagccgggaccggcggagcgcgagcgggccgggtgcgag
Cggactgaggagcggagcgcgactgccgggccgggcgaggcgggcggacggcggagagcg
Agggcgccgtcgccgtctctgagacgcacaaagggtcgaggctggggccgccgccgctc
agcgcgcgggcctggaaccggcagccccccggggctcggcgagaaggcggtgcgggcggcc
ATGGCCGGCTACGTGCCGGGTCAGCAGCCGGCCAACCGCAGCCAGGAGAAGGAGTTCGTG
 M  A  G  Y  V  P  G  Q  Q  P  A  N  R  S  Q  E  K  E  F  V
CAGGCGTACGAGGATGTGCTGGAGCGGTACAAAGgtagg
 Q  A  Y  E  D  V  L  E  R  Y  K
```

C. Exon 7 ORF prediction

```
atcccaggactggcttagctgcgtttttgctgagattaggagaggaaggaaATGGGAAAT
                                                  M  G  N
TCACTGGGCTGTTTTAAGGAGCCGAAAGAGTCAATAGCTATTCCTGAGAAGGCTCCCATA
 S  L  G  C  F  K  E  P  K  E  S  I  A  I  P  E  K  A  P  I
TCTCCTAAGAAAAGGGTTCGGTTCAAAAGGAGGTGGAGAGGCAAGAAAATCCCTACTCCA
 S  P  K  K  R  V  R  F  K  R  R  W  R  G  K  K  I  P  T  P
GAGGCATCTCACCAGGAAGAAACCTCAGAAGGAACTGGAGTCATTGAAGAGACTGAAACC
 E  A  S  H  Q  E  E  T  S  E  G  T  G  V  I  E  E  T  E  T
CTAACGAAGTTAACAGAGAGTCTCCAAAAGGAAGACGGAGTGGGAGGGGTAGAGCATACC
 L  T  K  L  T  E  S  L  Q  K  E  D  G  V  G  G  V  E  H  T
CCCCCAGATATTTTGCTGCCTGGGGACTCAGCCCCCAACTCACGGGTAGTCGATCGGGGG
 P  P  D  I  L  L  P  G  D  S  A  P  N  S  R  V  V  D  R  G
ATGATAGTACAGGTAAAGGAGAGATTCCAAGGGGAGGTCCAGACCGCCCACCTTTGTTA
 M  I  V  Q  V  K  E  R  F  Q  G  E  V  Q  T  A  H  L  L  L
GAGAATGAGTCATCAGTTGCTGGAGGGGTCTGGGATTCCCTGGAAGAGGGGATGACTGTC
 E  N  E  S  S  V  A  G  G  V  W  D  S  L  E  E  G  M  T  V
ATTGCTCACCTGCTTGATAACCCAGCAGAAAGGAACTGCGAGAAGTCAGTGAGCCAACTG
 I  A  H  L  L  D  N  P  A  E  R  N  C  E  K  S  V  S  Q  L
GTGGAATTTCCTAGGACAGCATCCTGCAGCAGCAGGGCTGTGTTGCTGCCTTTGCAAGGA
 V  E  F  P  R  T  A  S  C  S  S  R  A  V  L  L  P  L  Q  G
GAGACTGCAGTGGAGAAAGGAAATATTCAGCGTGGGTTTCGGAGCTGTGCTTTGCCTAGG
 E  T  A  V  E  K  G  N  I  Q  R  G  F  R  S  C  A  L  P  R
ACAGACTACCCCACTGATAAAGGAAATCAAGAACAATTTTCAGAGGGCTGGAGTGTGGAG
 T  D  Y  P  T  D  K  G  N  Q  E  Q  F  S  E  G  W  S  V  E
GAAGGAACCAAGAGTGTTTCAGGTGCCCCTCAGACAGCTTCCTGGATTATAGAATGTTCT
 E  G  T  K  S  V  S  G  A  P  Q  T  A  S  W  I  I  E  C  S
GTTTCTTCATTACTACTGGACCAGCCTGGAGGCCAAAGACGCACGGAGCCTTCCCATGTG
 V  S  S  L  L  L  D  Q  P  G  G  Q  R  R  T  E  P  S  H  V
GGTCAAGTGCCCCCCCAGGATTCCAGACTGCCTACTTCTCAGAGTGATTTGTCCATCAGT
 G  Q  V  P  P  Q  D  S  R  L  P  T  S  Q  S  D  L  S  I  S
GGTGTGACTGTGAGCATTTTGCCCTCCTCCTCTGGCTATGGCAGTGATGGGCCACACATA
 G  V  T  V  S  I  L  P  S  S  S  G  Y  G  S  D  G  P  H  I
CATGGGATCCAGCCTAAAGATACAGAACCTGAAAAGAGCTCTACTTCCTTCTCAGAAGAG
 H  G  I  Q  P  K  D  T  E  P  E  K  S  S  T  S  F  S  E  E
GATGGCACTCTTTCTCTGGAGGCAAGCCACACCCCATCATGGGGTCTGGAAGAGgtagg
 D  G  T  L  S  L  E  A  S  H  T  P  S  W  G  L  E  E
```

D. Exon 38 ORF prediction

```
ggctctgattcagttttctttttcttctctttcctttttttttttttttttaggaatgc
agagcagttgctggggtgattgacctaggcacagtggagatatttcccatcttcaaagcc
ATGCAAAAGGGCCTCCTTGACCAAGACACAGGCCTAGTGCTTCTGGAATCTCAGGTTATC
 M  Q  K  G  L  L  D  Q  D  T  G  L  V  L  L  E  S  Q  V  I
ATGTCTGGCCTCATTGCCCCTGAGACGGGTGAAAACCTCTCTTTGGAGGAGGGCATAGCC
 M  S  G  L  I  A  P  E  T  G  E  N  L  S  L  E  E  G  I  A
AGAAACCTCATTAATCCCCAGATGTACCAGCAGCTCCGGGAGCTACAGGATGCCCTGGCC
 R  N  L  I  N  P  Q  M  Y  Q  Q  L  R  E  L  Q  D  A  L  A
```

Fig. 3. Predicted 5′UTR and ORF of human MACF1 isoforms. **(A)** Exon 1 contains isoform 2-specific N-terminal sequence, based upon comparison with human MACF1-2 cDNA sequence. **(B)** Exon 2 contains isoform 1-specific N-terminal sequence, based on comparison with mouse MACF1-1 cDNA sequence (U67203). **(C)** Exon 7 contains isoform 3-specific N-terminal sequence, based on mouse MACF1-3 cDNA sequence (U67205). **(D)** Exon 38 contains isoform 4-specific N-terminal sequence; presented is a partial sequence of predicted N-terminus, including the first of 8 plectin repeats. Sequences of predicted ORFs are in upper case letters; predicted 5′ UTR sequences are in lower case. Putative initiation Met codons are presented in bold; upstream in-frame stop codons are underlined.

UTR riboprobe specific for MACF1 (corresponding to part of terminal exon 102) (Fig. 5A). This mouse probe had 90% sequence identity with the human MACF1 cDNA and would therefore cross-hybridize to human transcripts. Furthermore, this 3′UTR sequence had no significant sequence homology to any other mammalian sequences and thus was gene specific. MACF1 was expressed in all tissues examined, although the levels of expression varied. Hybridization was particularly strong in pituitary, adrenal, thyroid, salivary gland, mammary glands, pancreas, heart, and skeletal muscle. In a subsequent hybridization of the same human dot blot with a MACF1-4 (isoform 4)-specific riboprobe (corresponding to part of exon 38), signals were visible in all tissues, with the strongest hybridization signals in heart, lung, pituitary gland, and placenta (data not shown). This indicates differential tissue distribution among MACF1 transcripts. We also performed PCR surveys, using isoform-specific primers to determine expression patterns of various isoforms. Our survey of a collection of cDNA libraries prepared from 10 human tissues showed that MACF1-4-specific PCR product was detectable only in lung, heart, pituitary, and placenta, but was not visible in brain, kidney, liver, pancreas, skeletal muscle, or HepG2 cell line. These results were consistent with the fact that we were able to clone the MACF1-4-specific sequences from human heart and pituitary cDNA libraries (clones 2 and 3 in Fig. 1). On the other hand, MACF1-3 is expressed in lung, heart, placenta, as well as kidney and pituitary gland. Although the PCR results might depend on construction of individual cDNA libraries and might not be conclusive for transcript expression in any particular tissue, they appeared to reflect relative abundance of transcripts.

We determined the size and distribution of Macf1 transcripts by Northern blot analysis of total RNA from several mouse tissues, using a mouse riboprobe derived from the same mouse 3′UTR clone used in the human dot-blot experiment. Three large bands with estimated sizes of 12, 15, and 17 kb (Fig. 5B) were detected in several tissues, including brain, heart, ovary/uterus, kidney, and skeletal muscle. The relative intensities of these three bands varied among tissues; the 15-kb band was most abundant in the brain, whereas the 17-kb band was most abundant in the ovary/uterus and heart. As the probe used corresponds to the 3′UTR of Macf1 cDNAs, these three major bands reflected the presence of three (or more) Macf1 transcripts of various lengths that share the same 3′UTR. The largest 17-kb band is likely to represent the 19.2-kb MACF1-4 cDNA presented in this study, as well as the 17.7-kb Macf1-2 cDNA cloned previously. In a subsequent hybridization of the same Northern blot, a MACF1-2-specific riboprobe detected the 17-kb and 12-kb bands in brain and ovary/uterus; however, the MACF1-2 probe failed to detect either band in the heart. It is thus evident that some bands represent more than one transcript. It is not clear whether the 15-kb band corresponds to transcripts for isoforms 1 and/or 3 or additional isoforms that are yet to be identified.

Discussion

We sequenced BAC/PAC clones to determine the genome organization of the human MACF1 gene on Chr 1p32 between the NDUFS5 (telomeric) and BMP8/MYCL1 (centromeric) genes. Analysis of the genes for large transcripts like the MACF1 mRNA in either the public Human Genome Project or the Celera database indicated that neither group was successful in predicting the complete gene structure of MACF1 (called ACF7 in that analysis; Aach et al. 2001). First, in both the public and Celera databases the genomic region that contains the first 42 exons is currently in draft format with several gaps. Second, there are no ESTs in dbEST that define the structure of exon 38. In this study, we provide the first experimental evidence that this large genomic region actually codes for functional protein domains by isolating two human heart and pituitary gland cDNAs derived from this region.

A.

```
SR1     680  QSLHKFVSRATAELIWLNEKEEEELAYDWSD---NNSNISAKRNYFSELTMELEEKQ-DVFRSLQDTAELLSLENHPA--KQTVEAYSAAVQSQLQWMKQLCLCVEQHVKE      784
SR2    1455                                K---ESTDIEKAILEQQVLSEELTTKK-EQVSEAIKASQIFLAKHGHklsEKEKKQISEQLNALNKAYHDLCDGSANQLQQLQS  1534
SR3    1922  NQHTQLEGRLQDLRAWVGNKNLILNSKGSNSei-DVDSLNLCLQQYEDLKQPMAERK-AQLDALAFDIQFFISEHAQdlsPQQNRQMLRLLNELQRSFQDILEQTAAQVDALQG  2033
SR4    1547  KQQNTCHQQLEDLCSWVGQAERALAGHQGRTtqqDLSALQKNQSDLKDLQDDIQNRA-TSFATVVKDIEGFMEENQTklsPRELTALREKLHQAKEQYEALQEETRVAQKELEE  1659
SR5    1815  DELQKFLQDHKEFESWLERSEKELENMHKGGs--SPETLPSLLKRQGSFSEDVISHK-GDLRFVTISGQKVLDME-----NSFKE                                1891
SR6    1932  GQYHQFQNSADSLQAWMQACEANVEKLLSDTaasDPGVLQEQLATTKQLQEELAEHQ-VPVEKLQKVARDIMEIEGEp--APDHRHVQETTDSILSHFQSLSYSLAERSSLLQK  2042
SR7    2260  EQLDEFRKLVRTFQKWLKETEGSIPPTETSV---SAKELEKQIEHLKSLLDDWASKG-TLVEEINYKGTSLENLIME,,,CKDLTEIQCDMSDVNLKYEKLGGVLHERQESLQA  2395
SR8    2398  NRMEEVHKEANSVLQWLESKEEVLKSMDAMSspTKTETVKAQAESNKAFLAELEQNS-PKIQKVKEALAGLLVTYPN---SQEAENWKKIQEELNSRWERATEVTVARQRQLEE  2507
SR9    2525                WLMEKELMMGVLGPLSi--DPNMLNAQKQQVQFMLKEFEARR-QQHEQLNEAAQGILTGPGDv--SLSTSQVQKELQSINQKWVELTDKLNSRSSQIDQ  2618
SR10   2621  VKSTQYQELLQDLSEKVRAVGQRLSVQSAISt--QPEAVKQQLEETSEIRSDLEQLD-HEVKEAQTLCDELSVLIGE---QYLKDELKKRLETVALPLQGLEDLAADRINRLQA  2728
SR11   2731  ASTQQFQQMFDELRTWLDDKQSQQAKNCPISa--KLERLQSQLQENEEFQKSLNQHS-GSYEVIVAEGESLLLSVPP---GEEKRTLQNQLVELKNHWEELSKKTADRQSRLKD  2838
SR12   2841  QKAQKYQWHVEDLVPWIEDCKAKMSELRVTL---DPVQLESSLLRSKAMLNEVEKRR-SLLEILNSAADILINSSEA-----DEDGIRDEKAGINQNMDAVTEELQAKTGSLEE  2945
SR13   3187                      EAEALQWVVGTE----VEIINQQLADFKMFQKEqVDPLQMKLQQVNGLGQGLIQSAGK---DCDVQGLEHDMEEINARWNTLNKKVAQRIAQLQE  3274
SR14   3277  LHCGKFQDALEPLLSWLADTEELIANQKPPSa--EYKVVKAQIQEQKLLQRLLDDRK-ATVDMLQAEGGRIAQSAEL----ADREKITGQLESLESRWTELLSKAAARQKQLED  3383
SR15   3386  VLAKQFHETAEPISDFLSVTEKKLANSEPVGt--QTAKIQQQIIRHKALEEDIENHA-TDVHQAVKIGQSLSSLTS----PAEQGVLSEKIDSLQARYSEIQDRCCRKAALLDQ  3492
SR16   3495  SNARLFGEDEVEVLNWLAEVEDKLSSVFVKDf--KQDVLHRQHADHLALNEEIVNRK-KNVDQAIKNGQALLKQTT----GEEVLLIQEKLDGIKTRYADITVTSSKALRTLEQ  3601
SR17   3604  QLATKFQSTYEELTGWLREVEEELATSGGQSp--TGEQIPQFQQRQKELKKEVMEHR-LVLDTVNEVSRALLE                                          3673
SR18   3713  QRSQQVEQAADAELAWVAETKRKLMALGPIRl--EQDQTTAQLQVQKAFSIDIIRHK-DSMDELFSHRSEIFGTCG----EEQKTVLQEKTESLIQQYEAISLLNSERYARLER  3819
SR19   3832            EELSPWIEETRALIAQLPSPAi--DHEQLRQQQEEMRQLRESIAEHK-PHIDKLLKIGPQLKELN-----PEEGEMVEEKYQKAENMYAQIKEEVRQRALALDE  3927
SR20   3982                                                    ELEKLQPSFEALKRRGEELIGRSQGadk-DLAAKEIQDKLDQMVFFWEDIKARAEEREIKFLD  4043
SR21   4046  ELAEKFWYDMAALLTTIKDTQDIVHDLESPGi--DPSIIKQQVEAAETIKEETDGLH-EELEFIRILGADLIFACG----ETEKPEVRKSIDEMNNAWENLNKTWKERLEKLED  4152
SR22   4155  QAAVQYQDTLQAMFDWLDNTVIKLCTMPPVGt--DLNTVKDQLNEMKEFKVEVYQQQ-IEMEKLNHQGELMLKKATD---ETDRDIIREPLTELKHLWENLGEKIAHRQHKLEG  4262
SR23   4265  LALGQFQHALEELMSWLTHTEELLDAQRPISg--DPKVIEVELAKHHVLKNDVLAHQ-ATVETVNKAGNELLESSA----GDDASSLRSRLEAMNQCWESVLQKTEERFQQLQS  4371
SR24   4374  QQAQGFHSEIEDFLLELTRMESQLSASKPTGg--LPETAREQLDTHMELYSQLKAKE-ETYNQLLDKGRLMLLSRDD---SGSGSKTEQSVALLEQKWHVVSSKMEERKSKLEE  4481
SR25   4484  NLATEFQNSLQEFINWLTLAEQSLNIASPPSl--ILNTVLSQIEEHKVFANEVNAHR-DQIIELDQTGNQLKFLSQ----KQDVVLIKNLLVSVQSRWEKVVQRSIERGRSLDD  4590
SR26   4593  KRAKQFHEAWKKLIDWLEDAESHLDSELEISn--DPDKIKLQLSKHKEFQKTLGGKQ-PVYDTTIRTGRALKEKTLL---PEDSQKLDNFLGEVRDKWDTVCGKSVERQHKLEE  4700
SR27   4703  LFSGQFMDALQALVDWLYKVEPQLAEDQPVHg--DLDLVMNLMDAHKVFQKELGKRT-GTVQVLKRSGRELIENS-----RDDTTWVKGQLQELSTRWDTVCKLSVSKQSRLEQ  4808
SR28   4811  KQAEVFRDTVHMLLEWLSEAEQTLRFRGALPD--DTEALQSLIDTHKEFMKKVEEKRVDVNSAVAMGEVILAVCH-----PDCITTIKHWITIIRARFEEVLTWAKQHQQRLET  4916
SR29   4919  SELVANAELLEELLAWIQWAETTLIQRDQEPipqNIDRVKALIAEHQTFMEEMTRKQ-PDVDRVTKT                                                 4984

Consensus      lllqqFqrdadelesWisekeallssedygk...DlesvqaLlkkHealeaelaaheqdrvkqlnelaqkLieegea...hpdseeikerleeLnerWeaLlelaaeRrqkLee
               ═══════════ Helix A ═══ Turn ═══ Helix B ═══ Turn ═══ Helix C ═══
```

B.

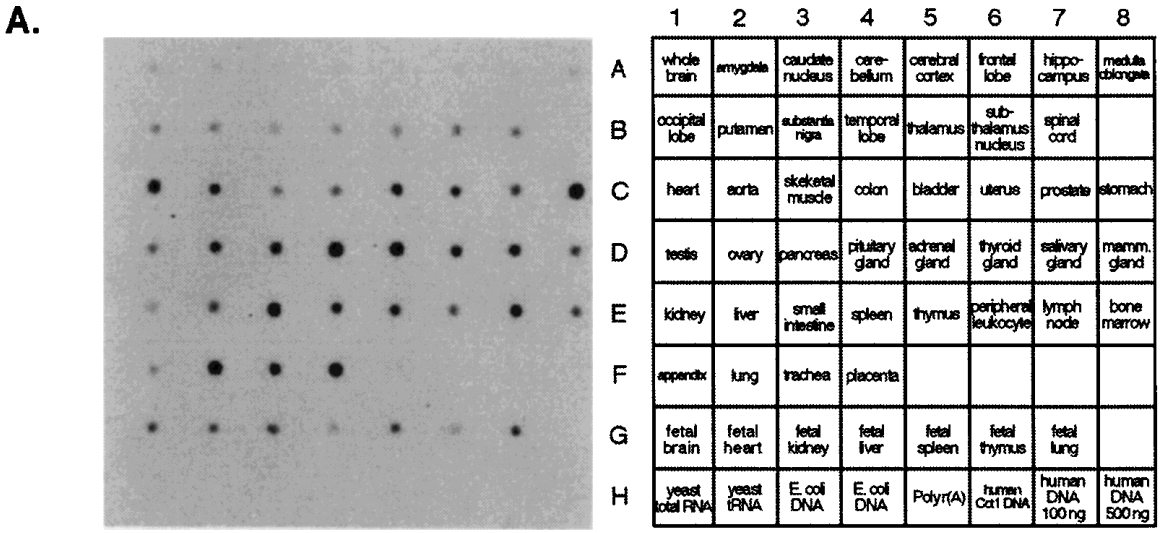
```
PR1     12   LVLLESQVIMSGLIAPETGENLSLEEGIARNLINPQMYQQLRELQ   56
PR2     89   LKILEAHLATGGFSLSPSENCINLEEAFHQGLISAWLHSVLESYL  133
PR3    203   VRLLEAQLFAGGIVDPRTGHRLTVEEAVRHNLIDQDMA-------  240
PR4    241   CAILIRQLQTGGIIDTVTGQRLTIDEAVSNDLVAAKIALVILESL  285
PR5    725   LNVLSAQLLDGGIFHEQTGQKLLLNEAISRGIVPSHTAVKLMEKL  769
PR6    801   HNVLMADKAISGVLDPRTQTLCSVKDAVTVGLLDKETATRILERQ  845
PR7    935   VRLLTKQVVDGGIIHHISGMRLSVDNAFRHGLIGEDLAEKLKRVE  979
PR8   1121   LKVLEAQANTGGIIDTATGKRLTLASALEEKLVDENMVRIIASHQ  1165

Consensus      qrlLeaqaatGGIiDPetgerLsvyeAlkrGLvdpetaqkLleae
```

Fig. 4. Sequence alignment of MACF1 spectrin and plectin repeats. **(A)** Alignment of 29 spectrin repeats encoded by the human MACF1 gene. Repeats were identified by Pfam analysis (E value < 0.05), aligned, and compared to the consensus sequence. The spectrin repeat is numbered based on the order of occurrence in the genomic sequence and listed in the left-most column (SR1-SR29). Coordinates of all repeats, except SR3, are based on the published MACF1-2 protein sequence (Macrophin, GenBank BAA83821); SR3 coordinates (italics) are from MACF1-4 sequence (AF317696). Gaps (−) are inserted to optimize the alignment with the consensus sequence. The (,) in SR7 denotes additional sequences encoded by exon 55, which is alternatively spliced and present in both the trabeculin-α (GenBank AAF06360) and MACF1-4 sequences, but absent in Macrophin. Secondary structure prediction of α-helices and turns are noted under the spectrin repeat consensus sequence. Proline (P) residues are in bold, indicating potential kinks in the secondary structure of the repeats. Vertical lines denote the position of introns. The intron position in helix B has been conserved, as seen previously among spectrin repeats in dystrophin (Koenig and Kunkel 1990). **(B)** Alignment of 8 plectin repeats (PR1–PR8) with the consensus sequence predicted by Pfam analysis (E value < 0.05). Plectin repeats are found only in MACF1-4, not in three other MACF1 isoforms. "Plectin repeats" are tandem repeats described first in plectin as a large "GC domain" (for the C-terminal globular domain) containing 6 long "GC repeats" of 245–343 amino acids each (Wiche et al. 1991). These GC repeats share a highly conserved central region that is constructed from nine tandem repeats of 19-residue motifs plus two flanking repeats. The Pfam database re-defines these tandem repeats and identifies them as "plectin repeats" with 45 amino acids in each unit (about three repeats of the same 19-residue core in each). These tandem repeats comprise the entire C-terminal globular domain of plectin: 20 copies of 45-aa plectin repeat.

MACF1 comprises at least 102 exons, spans over 270 kb, and gives rise to four major isoforms by alternative splicing. Exon 1 codes for isoform 2-specific 5′ UTR and N-terminal sequences, and exon 2 codes for isoform 1-specific N-terminal sequences. Exons 3–6 and 8–10 encode the two halves of the complete N-terminal actin-binding domain ABD that is present in both isoforms 1 and 2. An internal promoter preceding exon 7, which codes for isoform 3-specific sequences, bypasses the first six exons and results in isoform 3 lacking the first half of the ABD. Exons 38–42 are used only in isoform 4. In contrast to these three previously described isoforms, the newly identified isoform 4 (MACF1-4) does not appear to contain the ABD. Instead, MACF1-4 contains N-terminal plectin repeats that are encoded by a single large exon, exon 38, and is the only isoform known to contain plectin repeats. Following these plectin repeats and one spectrin repeat, the cDNA sequences are identical to the previously cloned MACF1-2 and code for the central α-helical rod-like domain composed of spectrin repeats, and the C-terminal globular domain that includes two EF hands, a Gas2 homology domain, and a serine/glycine-rich region.

The MACF1 gene appears to be a hybrid between genes encoding two different classes of cytoskeletal linker proteins, the plakin family and the spectrin/dystrophin family. Similarities to the gene for plectin, a well-characterized plakin, can be seen at both the 5′ and 3′ ends of MACF1. For example, the exon-intron boundaries for the N-terminal ABD in MACF1 are identical to those in plectin. Furthermore, one large exon encodes all the plectin repeats in MACF1-4, as in the plectin gene. The C-terminal end of MACF1, like the C-terminus of plectin, is serine/glycine-rich and contains GSR repeats. Similarity to the dystrophin gene is apparent in the genomic organization of the spectrin repeats. Each repeat is encoded by several small exons. The intron positions



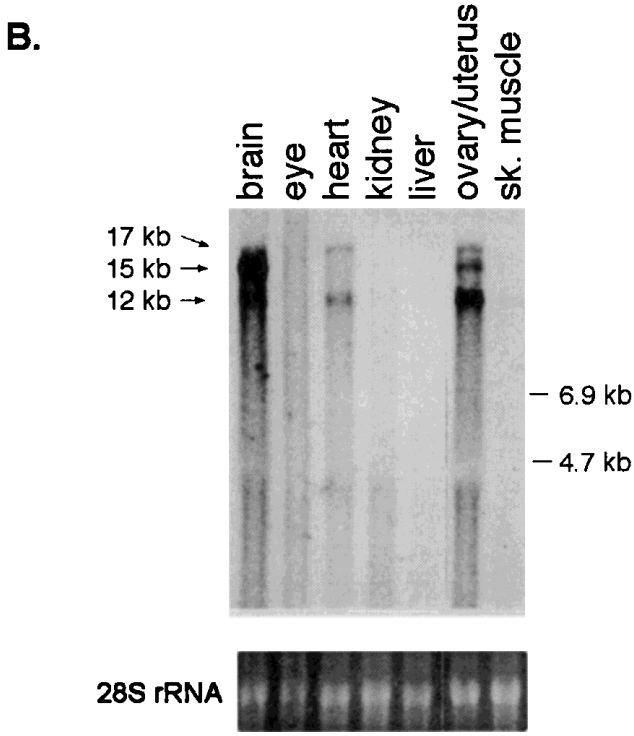


Fig. 5. Expression of MACF1. **(A)** Human RNA dot-blot analysis. Clontech human dot-blot containing polyA RNA from 50 tissues was hybridized with an [α-^{32}P]UTP-labeled anti-sense riboprobe derived from a mouse cDNA corresponding to the 3′UTR of MACF1-4 cDNA. **(B)** Northern blot analysis. Total RNA was isolated from various tissues of 4-week-old mice. Total RNA (15 μg each lane) was subjected to gel electrophoresis, transferred, and immobilized to Nytran membrane. **Top:** The membrane was hybridized with an anti-sense riboprobe derived from the same mouse 3′UTR clone as in the human dot blot analysis. Three major transcripts of 12, 15, and 17 kb were visible in brain, heart, kidney, ovary/uterus, and skeletal muscle. **Bottom:** Ethidium bromide-stained RNA gel before transfer to visualize the 28S ribosomal RNAs as a control for RNA loading.

within each spectrin repeat of MACF1 are highly conserved and very similar to those identified in the gene for dystrophin. Somewhat surprisingly, however, the exons encoding the plectin repeats are situated between the exons for the spectrin repeats in MACF1.

This hybrid gene structure appears to be an ancient, rather than a recent, event. Several genes identified in lower eukaryotic genomes have a structure similar to mammalian MACF1. The *Drosophila melanogaster* gene *shot*, the largest gene (>69 kb) in the *Drosophila* genome (Adams et al. 2000), encodes proteins given various names based on the phenotype of the mutant in which it was first identified: kakapo, groovin, and short stop. The alternative splicing seen at the 5′ end of MACF1 is also evident in the five isoforms of *Drosophila shot*. Two isoforms have an N-terminal ABD (Strumpf and Volk 1998), two have only a partial ABD, and a truncated isoform lacks the ABD entirely (Lee et al. 2000). There is a high degree of sequence conservation between shot/kakapo and mammalian MACF1 in the N-terminal ABD and C-terminal Gas2 region (74% similarity). The predicted secondary structure of the spectrin repeats in the central rod domain is conserved between the mammalian and *Drosophila* proteins, although the primary sequence is not. Genomic analysis revealed that *Drosophila shot* has larger exons and fewer introns. For example, the ABD is encoded by four rather than seven exons. Two large exons encode 21 of the 27 spectrin repeats in the central rod domain. Furthermore, the intron that was conserved in helix B of the mammalian spectrin repeat is not present in the *shot* gene.

Analysis of the *C. elegans* genome identified only one region with high sequence similarity to MACF1. GeneFinder predictions

currently suggest three consecutive ORFs (5′ to 3′: ZK1151.1, ZK1151.3, ZK1151.2a/b, GenBank Z93398); however, these ORFs may actually constitute a single gene, since the order of protein domains in the *C. elegans* ORFs is identical to that of human MACF1. The exon-intron boundaries in the first half of this conceptual *C. elegans* gene, particularly in the ABD, are also similar to those in MACF1. Experimental evidence is needed to confirm the genomic organization of the *C. elegans* gene and to determine whether or not alternative splicing occurs, as it does in the mammalian and *Drosophila* genes. Interestingly, the *Drosophila* and *C. elegans* genes each contain an exon for sequences with weak homology (E value > 0.10) to the plectin repeat (two to three repeats). This single large exon is positioned between exons for spectrin repeats, similar to that in mammalian MACF1. These common features in human MACF1, *Drosophila shot* and the *C. elegans* homolog suggest that these genes are derived from a common ancestral gene.

The plectin repeat, a sequence first described in plectin, is a distinctive feature of all plakins and is not found in other subfamilies of cytoskeletal linker proteins. In most plakins, these repeats are found at the C-terminus. MACF1-4 is the only MACF1 isoform shown to contain plectin repeats, which are located in the N-terminus and are preceded by spectrin repeats. While alternative splicing occurs in the 5′ end of plakins, such as plectin and dystonin/BPAG1n, to give rise to various N-termini, there are no known variants lacking plectin repeats, suggesting an important role of plectin repeats in the function of plakins. In plectin, these repeats harbor a domain required for intermediate filament binding (Nikolic et al. 1996), as well as binding site(s) for the β4 subunit of the hemidesmosomal α6β4 integrin receptor (Rezniczek et al. 1998). A similar functional role for the repeats in other plakins has yet to be demonstrated. Since plectin repeats are the common domain feature among intermediate filament-interacting plakins, we propose that plectin repeats may provide the key structure for interaction with intermediate filaments. If it were the case, MACF1-4 would provide a link to intermediate filaments via its N-terminal plectin repeat domain, while MACF1-2 links to microfilaments via its N-terminal actin binding domain (Leung et al. 1999; Sun et al. 1999). Such sequence variation in the N-termini among different isoforms is likely to enhance functional diversity of MACF1 gene products.

In contrast to the variable domain structure at the N-termini, the central rod domain of MACF1 is invariant and is composed mainly of spectrin repeats. This conclusion is based on the published cDNA sequence of MACF1 isoform 2 and the data on isoform 4 presented in this report. Similar to plectin repeats, spectrin repeats are thought to arise from internal duplications. Spectrin repeats are 109 residue motifs that are also found in several cytoskeletal linker proteins, such as α-actinin, spectrin, dystrophin, and utrophin. The spectrin repeat forms three α-helices (A, B, and C) separated by proline-rich linkers (Grum et al. 1999; Yan et al. 1993). The intron position in helix B is conserved in most spectrin repeats of MACF1, a feature remarkably similar to dystrophin (Koenig and Kunkel 1990; Winder et al. 1995).

The Gas2 homology domain, a distinctive feature of MACF1 and its related large cytoskeletal linker proteins, has 66% similarity (or 38% identity) to a short, C-terminal region (amino acids 238–271) of mouse Gas2 (growth arrest-specific 2) protein (Cowled et al. 1994; Manzow et al. 1996). Gas2 is a component of the actin microfilament network (Brancolini et al. 1992) and a substrate of Caspases 3 and 7 in the apoptosis process (Sgorbissa et al. 1999). A similar protein is encoded by human GAR22 (Gas2-related sequence on human Chr 22) (Zucman-Rossi et al. 1996). In MACF1, the Gas2 domain, together with the adjacent serine/glycine-rich region, binds to and appears to stabilize microtubules (Sun et al. 2001).

Over-expression of either a full length or a truncated Gas2 protein with a C-terminal deletion induces cytoskeletal rearrangement and changes in cell shape that mimic the effects of proteolytic cleavage of endogenous Gas2 in apoptotic cells. The domain that is conserved in MACF1 (the Gas2 homology domain) does not appear to be required for cytoskeletal rearrangement, since overexpression of a C-terminal deletion construct that removes the domain still leads to cytoskeletal rearrangements. Deletion of 36 additional amino acids from the Gas2 protein, however, completely abolishes these morphological changes (Brancolini et al. 1995). The domain critical in inducing cytoskeletal changes is thus likely to reside in these 36 amino acids.

The genome of *C. elegans* contains two genes with Gas2 homology domains: the MACF1 homolog (ZK1151) and a much smaller gene (D2096.11) for a hypothetical protein of 1021 amino acids. This second hypothetical protein consists of an N-terminal ABD and two copies of the Gas2 domain in the C-terminal half, but lacks intervening spectrin or plectin repeats. *Drosophila* genome also contains two genes with Gas2 domains: *shot* (MACF1 homolog) and a gene similar to *Gas2*/GAR22 (CG3973, GenBank AE003438). In mammals, only one additional gene besides MACF1, *Gas2,* and GAR22 is known to contain Gas2-homology sequences. This expressed gene, MACF2, is located on human Chr 6p12, near BPAG1, and has also been called KIAA0728 (Nagase et al. 1998) or trabeculin-β (Sun et al. 1999). The KIAA0728 partial cDNA encodes only 1065 amino acids of MACF2, including the last five spectrin repeats, two EF hands, the Gas2 domain, and the serine/glycine-rich region. MACF2 shares 80% sequence homology with MACF1 throughout the coding region that has been cloned thus far. In addition, the intron/exon positions of the sequenced region of MACF2 are identical to those in MACF1 (data not shown). MACF1 and MACF2 are therefore likely to be paralagous genes. On the other hand, MACF1 on human Chr 1p32 and the previously cloned mouse *Macf/Acf7* are orthologous genes based on the conserved synteny between the human and the mouse genomes in these two chromosomal regions, as well as the high sequence identity (90%) in the 3′UTRs.

We observed differential expression patterns of three large MACF1 transcripts among tissues. The complex alternative splicing and the use of tissue-specific internal promoters seen in MACF1 are common features of genes for other cytoskeletal linker proteins, such as dystonin/BPAG1, plectin, and dystrophin. Dystonin/BPAG1 gene codes for at least four isoforms. Isoforms 1 and 2 have the full-sized ABD, and isoform 3 has only the second half of the ABD, whereas isoform 4 uses an internal promoter and has no ABD at all. BPAG1 isoforms are differentially expressed among tissues; isoforms 1–3 (BPAG1n1, 2, and 3) are expressed predominantly in neuronal tissues, whereas isoform 4 (BPAG1e) is controlled by a keratinocyte-specific promoter (Brown et al. 1995). Plectin has at least 12 transcripts with different 5′ ends as a result of alternative splicing of 13 coding or non-coding exons; expression of transcripts is also tissue specific (Fuchs et al. 1999). The Dystrophin gene encodes for at least 15 isoforms of various lengths containing different segments of the basic dystrophin sequence by use of tissue-specific promoters (Amalfitano et al. 1997). Such complex alternative splicing results in multiple isoforms that may reflect the diverse function of these cytolinkers. For example, plectin has been shown to interact with intermediate filaments, actin microfilaments, microtubule network via interaction with microtubule-associated proteins, as well as with the nuclear lamina, various components in desmosomes and hemidesmosomes, and signaling molecules (for reviews, see also Steinbock and Wiche 1999). More recently, plectin has been shown to be a substrate of the caspase cascade (Stegh et al. 2000).

The different MACF1 isoforms reflect functional diversity of these large cytoskeletal linkers. Several studies on mammalian MACF1-2 have assigned functions to the N- and C-terminal domains. MACF1-2 has been postulated to crosslink the actin microfilament and microtubule networks. The N-terminal domain of MACF1-2 has been shown to bind actin, whereas the C-terminus of MACF1-2 associates with microtubules, probably via the Gas2

domain, as well as the serine/glycine-rich region (Karakesisoglou et al. 2000; Leung et al. 1999; Sun et al. 2001). The co-alignment of actin, microtubules, and MACF1 near intercellular junctions, where microtubules and microfilaments associate, provides additional support for this crosslinking function. A similar functional role has been proposed for *Drosophila* Kakapo, which is located at the termini of microtubule bundles and may provide links between the microtubule and the cortical actin networks (Strumpf and Volk 1998). Absence of kakapo causes detachment of these microtubule bundles from the basal membrane (Gregory and Brown 1998). On the other hand, MACF1-3, which lacks the first part of the N-terminal ABD, does not appear to associate with the actin network. This conclusion is based on studies of BPAG1n3, a neuronal isoform of dystonin/BPAG1, which associates with microtubules through an "M1 domain" immediately following the ABD (Karakesisoglou et al. 2000; Yang et al. 1999). Both BPAG1n3 and MACF1-3 have similar ABDs, and the "M1 domain" is present in both molecules.

What is the function of MACF1 isoform 4, which lacks the entire actin-binding domain and the M1 domain? Previous studies of MACF1-2 do not provide any clear evidence for MACF1-4 function, since the probes and antibodies used in those studies might not have detected this new isoform. It is expected that MACF1-4 interacts with molecules different from those of cytoskeletal proteins with ABD. If plectin repeats are indeed the keys for intermediate filament binding, MACF1-4 will provide a link between intermediate filaments (via the N-terminus) and microtubules (via the C-terminus). Even though MACF1-4 lacks the N-terminal ABD, an association with the actin network cannot be excluded, since truncated isoforms of dystrophin without an ABD still bind actin. In dystrophin, a cluster of basic spectrin repeats (repeats 11–14) is known to bind to F-actin via electrostatic interactions (Amann et al. 1998), although a similar basic cluster in utrophin does not bind to F-actin (Amann et al. 1999). Since alternative splicing is evident not only in the N-terminus but also in other parts of the protein, these minor variations may provide mechanisms for further fine tuning of functions such as binding affinity. Differential expression pattern of MACF1 transcripts among tissues, as seen by this and another recent study (Bernier et al. 2000), reflects the existence of multiple isoforms within any given tissues, or even cells. Experimental evidence is needed to clarify functional differences among these isoforms.

There are currently no mammalian genetic models associated with defects in MACF1. The MACF1 gene is located on human Chr 1p32 in the region of a dominant deafness locus, DFNA2. Mutations in Cx31 (Connexin31) and KCNQ4 that are centromeric to MACF1 have been shown to be responsible for deafness in some DFNA2 families (Coucke et al. 1999; Kubisch et al. 1999; Xia et al. 1998). However, no mutation in either gene has been found in the original Indonesian family (Coucke et al. 1999) that defined the DFNA2 locus. Since the MACF1 gene is located within the recombinant region defined by the Indonesian DNFA2 family, it remains a candidate gene for this or other deafness families mapped to this region. It is not yet clear which isoforms are present in the ear, nor what their functional roles may be. Sound conductance heavily depends on the integrity of delicate cytoskeletal networks and the stability of the sensory epithelium that is constantly challenged by acoustic stimulation. Defects in structural proteins, e.g., cadherins (Alagramam et al. 2001; Bolz et al. 2001; Bork et al. 2001; Di Palma et al. 2001; Wada et al. 2001; Wilson et al. 2001), collagens (Lemmink et al. 1997; Verstreken et al. 1996), connexins (Carrasquillo et al. 1997; Denoyelle et al. 1997; Kelley et al. 1998; Kelsell et al. 1997; Scott et al. 1998), and tectorins (Hughes et al. 1998; Verhoeven et al. 1998), have been shown to lead to hereditary deafness. A remaining challenge is to determine whether mutations in MACF1 do indeed lead to deterioration of auditory function. The complete structure of the MACF1 gene should assist these efforts.

Acknowledgments. This work was supported by NIH grants R01 DC02492 and P01 DC02982, and a grant from the National Organization for Hearing Research (M.I. Lomax). The authors thank Robert S. Winnicki for excellent technical support, Drs. D. Kohrman and J. Chamberlain for their critical comments on the manuscript, Lucy Rowe and Mary Barter (The Jackson Laboratory, Bar Harbor, Me.) for assistance with data analysis, and Dr. T. Glaser, Howard Hughes Medical Institute and Dept. of Human Genetics, for RNA from EBV-transformed human lymphoblastoid cells.

References

Aach J, Bulyk ML, Church GM, Comander J, Derti A et al. (2001) Computational comparison of two draft sequences of the human genome. Nature 409, 856–859

Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD et al. (2000) The genome sequence of *Drosophila melanogaster*. Science 287, 2185–2195

Alagramam KN, Murcia CL, Kwon HY, Pawlowski KS, Wright CG et al. (2001) The mouse Ames waltzer hearing-loss mutant is caused by mutation of Pcdh15, a novel protocadherin gene. Nat Genet 27, 99–102

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25, 3389–3402

Amalfitano A, Rafael JA, Chamberlain JS (1997) Dystrophin gene structure and mutation. In *Dystrophin: Gene, Protein and Cell Biology*, J Lucy, S Brown, eds. (Cambridge, UK: Cambridge University Press), pp 1–26

Amann KJ, Renley BA, Ervasti JM (1998) A cluster of basic repeats in the dystrophin rod domain binds F-actin through an electrostatic interaction. J Biol Chem 273, 28419–28423

Amann KJ, Guo AW, Ervasti JM (1999) Utrophin lacks the rod domain actin binding activity of dystrophin. J Biol Chem 274, 35375–35380

Bernier G, Mathieu M, De Repentigny Y, Vidal SM, Kothary R (1996) Cloning and characterization of mouse ACF7, a novel member of the dystonin subfamily of actin binding proteins. Genomics 38, 19–29

Bernier G, Pool M, Kilcup M, Alfoldi J, De Repentigny Y et al. (2000) Acf7 (MACF) is an actin and microtubule linker protein whose expression predominates in neural, muscle, and lung development [In Process Citation]. Dev Dyn 219, 216–225

Bolz H, von Brederlow B, Ramirez A, Bryda EC, Kutsche K et al. (2001) Mutation of CDH23, encoding a new member of the cadherin gene family, causes Usher syndrome type 1D. Nat Genet 27, 108–112

Bork JM, Peters LM, Riazuddin S, Bernstein SL, Ahmed ZM et al. (2001) Usher syndrome 1D and nonsyndromic autosomal recessive deafness DFNB12 are caused by allelic mutations of the novel cadherin-like gene CDH23. Am J Hum Genet 68, 26–37

Brancolini C, Bottega S, Schneider C (1992) Gas2, a growth arrest-specific protein, is a component of the microfilament network system. J Cell Biol 117, 1251–1261

Brancolini C, Benedetti M, Schneider C (1995) Microfilament reorganization during apoptosis: the role of Gas2, a possible substrate for ICE-like proteases. EMBO J 14, 5179–5190

Brown A, Bernier G, Mathieu M, Rossant J, Kothary R (1995) The mouse dystonia musculorum gene is a neural isoform of bullous pemphigoid antigen 1. Nat Genet 10, 301–306

Byers TJ, Beggs AH, McNally EM, Kunkel LM (1995) Novel actin crosslinker superfamily member identified by a two step degenerate PCR procedure. FEBS Lett 368, 500–504

Carrasquillo MM, Zlotogora J, Barges S, Chakravarti A (1997) Two different connexin 26 mutations in an inbred kindred segregating nonsyndromic recessive deafness: implications for genetic studies in isolated populations. Hum Mol Genet 6, 2163–2172

Collavin L, Buzzai M, Saccone S, Bernard L, Federico C et al. (1998) cDNA characterization and chromosome mapping of the human GAS2 gene. Genomics 48, 265–269

Coucke PJ, Van Hauwe P, Kelley PM, Kunst H, Schatteman I et al. (1999) Mutations in the KCNQ4 gene are responsible for autosomal dominant deafness in four DFNA2 families. Hum Mol Genet 8, 1321–1328

Cowled PA, Ciccarelli C, Coccia E, Philipson L, Sorrentino V (1994) Expression of growth arrest-specific (gas) genes in senescent murine cells. Exp Cell Res 211, 197–202

Dalpe G, Mathieu M, Comtois A, Zhu E, Wasiak S et al. (1999) Dystonin-deficient mice exhibit an intrinsic muscle weakness and an instability of skeletal muscle cytoarchitecture. Dev Biol 210, 367–380

Denoyelle F, Weil D, Maw MA, Wilcox SA, Lench NJ et al. (1997)

Prelingual deafness: high prevalence of a 30delG mutation in the connexin 26 gene. Hum Mol Genet 6, 2173–2177
Di Palma F, Holme RH, Bryda EC, Belyantseva IA, Pellegrino R et al. (2001) Mutations in Cdh23, encoding a new type of cadherin, cause stereocilia disorganization in waltzer, the mouse model for Usher syndrome type 1D. Nat Genet 27, 103–107
Dowling J, Yang Y, Wollmann R, Reichardt LF, Fuchs E (1997) Developmental expression of BPAG1-n: insights into the spastic ataxia and gross neurologic degeneration in dystonia musculorum mice. Dev Biol 187, 131–142
Fuchs E, Yang Y, Dowling J, Kouklis P, Smith E et al. (1997) Intermediate filament linker proteins. Soc Gen Physiol Ser 52, 141–148
Fuchs P, Zorer M, Rezniczek GA, Spazierer D, Oehler S et al. (1999) Unusual 5′ transcript complexity of plectin isoforms: novel tissue-specific exons modulate actin binding activity. Hum Mol Genet 8, 2461–2472
Gong TW, Hegeman AD, Shin JJ, Adler HJ, Raphael Y et al. (1996) Identification of genes expressed after noise exposure in the chick basilar papilla. Hear Res 96, 20–32
Gong TW, Winnicki RS, Kohrman DC, Lomax MI (1999) A novel mouse kinesin of the UNC-104/KIF1 subfamily encoded by the Kif1b gene. Gene 239, 117–127
Gregory SL, Brown NH (1998) Kakapo, a gene required for adhesion between and within cell layers in Drosophila, encodes a large cytoskeletal linker protein related to plectin and dystrophin. J Cell Biol 143, 1271–1282
Grum VL, Li D, MacDonald R, Mondragon A (1999) Structures of two repeats of spectrin suggest models of flexibility. Cell 98, 523–535
Guo L, Degenstein L, Dowling J, Yu QC, Wollmann R et al. (1995) Gene targeting of BPAG1: abnormalities in mechanical strength and cell migration in stratified epithelia and neurologic degeneration. Cell 81, 233–243
Hughes DC, Legan PK, Steel KP, Richardson GP (1998) Mapping of the alpha-tectorin gene (TECTA) to mouse chromosome 9 and human chromosome 11: a candidate for human autosomal dominant nonsyndromic deafness. Genomics 48, 46–51
Karakesisoglou I, Yang Y, Fuchs E (2000) An epidermal plakin that integrates actin and microtubule networks at cellular junctions. J Cell Biol 149, 195–208
Kelley PM, Harris DJ, Comer BC, Askew JW, Fowler T et al. (1998) Novel mutations in the connexin 26 gene (GJB2) that cause autosomal recessive (DFNB1) hearing loss. Am J Hum Genet 62, 792–799
Kelsell DP, Dunlop J, Stevens HP, Lench NJ, Liang JN et al. (1997) Connexin 26 mutations in hereditary non-syndromic sensorineural deafness. Nature 387, 80–83
Koenig M, Kunkel L (1990) Detailed analysis of the repeat domain of dytrophin reveals four potential hinge segments that may confer flexibility. J Biol Chem 266, 4560–4566
Kozak M (1987) An analysis of 5′-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res 15, 8125–8148
Kubisch C, Schroeder BC, Friedrich T, Lutjohann B, El-Amraoui A et al. (1999) KCNQ4, a novel potassium channel expressed in sensory outer hair cells, is mutated in dominant deafness. Cell 96, 437–446
Lee S, Harris KL, Whitington PM, Kolodziej PA (2000) Short stop is allelic to kakapo, and encodes rod-like cytoskeletal-associated proteins required for axon extension. J Neurosci 20, 1096–1108
Lemmink HH, Schroder CH, Monnens LA, Smeets HJ (1997) The clinical spectrum of type IV collagen mutations. Hum Mutat 9, 477–499
Leung CL, Sun D, Zheng M, Knowles DR, Liem RK (1999) Microtubule actin cross-linking factor (MACF): a hybrid of dystonin and dystrophin that can interact with the actin and microtubule cytoskeletons. J Cell Biol 147, 1275–1286
Liu CG, Maercker C, Castanon MJ, Hauptmann R, Wiche G (1996) Human plectin: organization of the gene, sequence analysis, and chromosome localization (8q24). Proc Natl Acad Sci USA 93, 4278–4283
Mahoney MG, Aho S, Uitto J, Stanley JR (1998) The members of the plakin family of proteins recognized by paraneoplastic pemphigus antibodies include periplakin. J Invest Dermatol 111, 308–313
Manzow S, Brancolini C, Marks F, Richter KH (1996) Expression of growth arrest-specific (Gas) genes in murine keratinocytes: Gas2 is specifically regulated. Exp Cell Res 224, 200–203
McLean WH, Pulkkinen L, Smith FJ, Rugg EL, Lane EB et al. (1996) Loss of plectin causes epidermolysis bullosa with muscular dystrophy: cDNA cloning and genomic organization. Genes Dev 10, 1724–1735
Nagase T, Ishikawa K, Suyama M, Kikuno R, Miyajima N et al. (1998) Prediction of the coding sequences of unidentified human genes. XI. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro. DNA Res 5, 277–286
Nikolic B, Mac Nulty E, Mir B, Wiche G (1996) Basic amino acid residue cluster within nuclear targeting sequence motif is essential for cytoplasmic plectin-vimentin network junctions. J Cell Biol 134, 1455–1467
Okuda T, Matsuda S, Nakatsugawa S, Ichigotani Y, Iwahashi N et al. (1999) Molecular cloning of macrophin, a human homolog of Drosophila kakapo with a close structural similarity to plectin and dystrophin. Biochem Biophys Res Commun 264, 568–574
Rezniczek GA, de Pereda JM, Reipert S, Wiche G (1998) Linking integrin alpha6beta4-based cell adhesion to the intermediate filament cytoskeleton: direct interaction between the beta4 subunit and plectin at multiple molecular sites. J Cell Biol 141, 209–225
Rowe LB, Nadeau JH, Turner R, Frankel WN, Letts VA et al. (1994) Maps from two interspecific backcross DNA panels available as a community genetic mapping resource. Mamm Genome 5, 253–274
Ruhrberg C, Watt FM (1997) The plakin family: versatile organizers of cytoskeletal architecture. Curr Opin Genet Dev 7, 392–397
Ruhrberg C, Hajibagheri MA, Parry DA, Watt FM (1997) Periplakin, a novel component of cornified envelopes and desmosomes that belongs to the plakin family and forms complexes with envoplakin. J Cell Biol 139, 1835–1849
Scott DA, Kraft ML, Stone EM, Sheffield VC, Smith RJH (1998) Connexin mutations and hearing loss. Nature 391, 32
Sgorbissa A, Benetti R, Marzinotto S, Schneider C, Brancolini C (1999) Caspase-3 and caspase-7 but not caspase-6 cleave Gas2 in vitro: implications for microfilament reorganization during apoptosis. J Cell Sci 112, 4475–4482
Stegh AH, Herrmann H, Lampel S, Weisenberger D, Andra K et al. (2000) Identification of the cytolinker plectin as a major early in vivo substrate for caspase 8 during CD95- and tumor necrosis factor receptor-mediated apoptosis. Mol Cell Biol 20, 5665–5679
Steinbock FA, Wiche G (1999) Plectin: a cytolinker by design. J Biol Chem 380, 151–158
Strumpf D, Volk T (1998) Kakapo, a novel cytoskeletal-associated protein is essential for the restricted localization of the neuregulin-like factor, vein, at the muscle-tendon junction site. J Cell Biol 143, 1259–1270
Sun D, Leung CL, Liem RK (2001) Characterization of the microtubule binding domain of microtubule actin crosslinking factor (MACF): identification of a novel group of microtubule associated proteins. J Cell Sci 114, 161–172
Sun Y, Zhang J, Kraeft SK, Auclair D, Chang MS et al. (1999) Molecular cloning and characterization of human trabeculin-alpha, a giant protein defining a new family of actin-binding proteins. J Biol Chem 274, 33522–33530
Verhoeven K, Van Laer L, Kirschofer K, Legan PK, Hughes DC et al. (1998) Mutations in the human α-tectorin gene cause autosomal dominant non-syndromic hearing impairment. Nat Genet 19, 60–62
Verstreken M, Claes J, Van de Heyning PH (1996) Osteogenesis imperfecta and hearing loss. Acta Oto-rhino-laryngol Belg 50, 91–98
Wada T, Wakabayashi Y, Takahashi S, Ushiki T, Kikkawa Y et al. (2001) A point mutation in a cadherin gene, Cdh23, causes deafness in a novel mutant, Waltzer mouse niigata. Biochem Biophys Res Commun 283, 113–117
Wiche G, Becker B, Luber K, Weitzer G, Castanon MJ et al. (1991) Cloning and sequencing of rat plectin indicates a 466-kD polypeptide chain with a three-domain structure based on a central alpha-helical coiled coil. J Cell Biol 114, 83–99
Wilson SM, Householder DB, Coppola V, Tessarollo L, Fritzsch B et al. (2001) Mutations in Cdh23 cause nonsyndromic hearing loss in waltzer mice. Genomics 74, 228–233
Winder SJ, Gibson TJ, Kendrick-Jones J (1995) Dystrophin and utrophin: the missing links! FEBS Lett 369, 27–33
Xia JH, Liu CY, Tang BS, Pan Q, Huang L et al. (1998) Mutations in the gene encoding gap junction protein beta-3 associated with autosomal dominant hearing impairment. Nat Genet 20, 370–373
Yan Y, Winograd E, Viel A, Cronin T, Harrison SC et al. (1993) Crystal structure of the repetitive segments of spectrin. Science 262, 2027–2030
Yang Y, Bauer C, Strasser G, Wollman R, Julien JP et al. (1999) Integrators of the cytoskeleton that stabilize microtubules. Cell 98, 229–238
Zucman-Rossi J, Legoix P, Thomas G (1996) Identification of new members of the Gas2 and Ras families in the 22q12 chromosome region. Genomics 38, 247–254