Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochemical and Biophysical Research Communications 309 (2003) 143–154
BBRC
www.elsevier.com/locate/ybbrc

A novel giant gene CSMD3 encoding a protein with CUB and sushi multiple domains: a candidate gene for benign adult familial myoclonic epilepsy on human chromosome 8q23.3–q24.1

Atsushi Shimizu,[a] Shuichi Asakawa,[a] Takashi Sasaki,[a] Satoru Yamazaki,[a] Hidehisa Yamagata,[b] Jun Kudoh,[a] Shinsei Minoshima,[a,1] Ikuko Kondo,[b] and Nobuyoshi Shimizu[a,*]

[a] *Department of Molecular Biology, Keio University School of Medicine, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan*
[b] *Department of Hygiene, Ehime University School of Medicine, Shitsukawa, Shigenobu, Onsen-gun, Ehime 781-0295, Japan*

Received 22 July 2003

Abstract

We identified a novel giant gene encoding a transmembrane protein with CUB and sushi multiple domains on the human chromosome 8q23.3–q24.1 in which benign adult familial myoclonic epilepsy type 1 (BAFME1/FAME, OMIM:601068) has been mapped. This giant gene consists of 73 exons and spans over 1.2 Mb on the genomic DNA region. It showed significant homology to two genes, CSMD1 gene on 8p23 and CSMD2 gene on 1p34, at reduced amino acid sequence level and hence we designated as CSMD3. The CSMD3 gene was expressed mainly in adult and fetal brains. We performed mutation analysis on the CSMD3 gene for seven patients with BAFME1/FAME, but no mutation was found in the coding sequence of the CSMD3 gene. Comparative genomic analysis revealed a conserved family of CSMD genes in the mouse and fugu genomes. Possible functions of the CSMD gene family are discussed.

Keywords: BAFME1; FAME; CUB; Sushi; CSMD; Epilepsy; Brain; Comparative genomics; Chromosome 8

We have constructed a BAC contig of 30 Mb corresponding to the human chromosome 8q22–q24.1, and performed a genomic DNA sequencing followed by extensive computer-aided gene annotation in combination with manual and experimental examination. We identified 120 genes, among which we focused on a giant gene (named CSMD3) of 1.2 Mb consisting of 73 exons. We determined the complete structure of CSMD3 gene and found that it encodes a transmembrane protein of CUB and sushi multiple domains and it is expressed mainly in fetal and adult brains, suggesting a good candidate for the pathogenic gene for the benign adult familial myoclonic epilepsy type 1 (BAFME1/FAME, OMIM:601068) which has been mapped to this chromosomal region.

The BAFME1/FAME is characterized by adult-onset tremulous finger movement, myoclonus, epileptic seizures, and a non-progressive course [1]. The age of onset of the disorder is between 18 and 50 years old. The BAFME1/FAME showed a clear autosomal dominant pattern of inheritance. Linkage analysis of Japanese patients indicated that the responsible gene for the BAFME1/FAME is located to chromosome 8q23.3–q24.1 [2,3]. Mikami et al. [3] showed a maximum two-point LOD score of 4.31 for a marker D8S555 with a recombination fraction of zero and a maximum multi-point LOD score of 5.42 for the interval between D8S555 and D8S1779 for a large Japanese pedigree. In this family, no recombination was observed with three markers, D8S1830, D8S555, and D8S1779, which are

* Corresponding author. Fax: 81-3-3351-2370.
E-mail address: shimizu@dmb.med.keio.ac.jp (N. Shimizu).
[1] Present address: Photon Medical Research Center, Hamamatsu University School of Medicine, 1-20-1 Handayama, Hamamatsu, Shizuoka 431-3192, Japan.

0006-291X/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/S0006-291X(03)01555-9

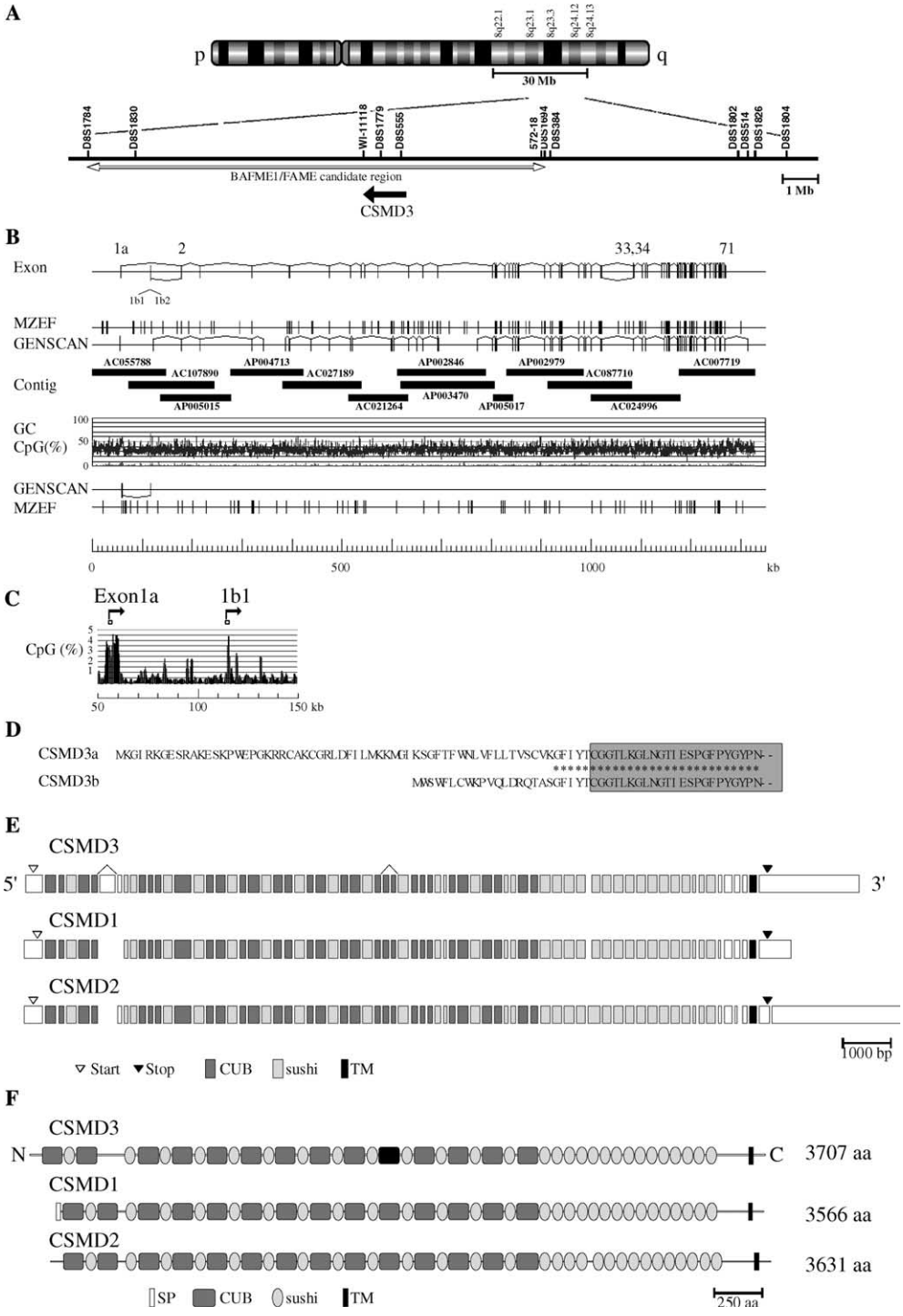

Fig. 1. A novel human gene CSMD3: chromosomal location, gene structure, CpG islands, and protein domains. (A) Diagram of chromosome 8, STS makers linked to BAFME1/FAME and location of CSMD3 gene. White arrow indicates BAFME1/FAME candidate region. Black arrow indicates 5′ to 3′ direction of the gene. (B) The exon–intron structure of the CSMD3 gene, positions of exons on the upper DNA strand predicted by GENSCAN and MZEF, BAC contig and accession numbers of BAC clone sequence data, GC/CpG contents, positions of predicted exons on the lower DNA strand. (C) Location of CpG islands and 5′ ends of two first exons. (D) N-terminal amino acid sequences of proteins translated from two transcripts (CSMD3a and CSMD3b). A part of CUB domain is shown by dark gray. (E) Exon structures of CSMD transcripts. The positions of start and stop codons are indicated by open and solid arrowheads, respectively. Exons encoding CUB, sushi, and transmembrane domains (TM) are depicted by dark gray, light gray and filled boxes, respectively. The connected lines above the transcript indicate alternative splicing of exons 7, 33, and 34. CSMD3 represents the transcript started at exon 1a. (F) Predicted structure of the CSMD proteins. Rectangles, squares, ovals, and vertical hatches represent putative signal peptides, CUB domains, sushi domains, and putative transmembrane domains, respectively. The blackened CUB domain is subject to skipping by alternative splicing. CSMD3 represents the protein translated from CSMD3a.

located within an 8-cM interval between D8S1784 and D8S1694, indicating that this 8-cM region contains a causative gene for BAFME1. For the linkage analysis of FAME families by Plaster et al. [2], we reviewed their data according to precise order of microsatellite markers based on the finished human genome sequence. We then found that the FAME locus is mapped on a region proximal to D8S514, which is overlapped with BAFME1 candidate region. According to these information, we focused on the 8-cM interval between D8S1784 and D8S1694 as target region of BAFME1/FAME. The mapped locus is distinct from three other epilepsy-related loci on chromosome 8q (KCNQ3, ECA1, and FEB1) [4–6].

Here, we report a novel giant gene CSMD3 of 1.2 Mb encoding a transmembrane protein and describe initial characterization of the gene in terms of its genomic structure, transcript variants, tissue expression profiles, evolutionary relationship with the CSMD family genes in a phylogenetic tree, and mutation analysis as a candidate gene for a type of epilepsy BAFME1/FAME.

Materials and methods

BAFME1/FAME patients. The DNA samples were obtained from healthy individuals and patients with BAFME1/FAME after receiving informed consent for this study [2,3].

Construction of BAC contigs and sequencing. BAC contig of 30 Mb covering 8q22.1–q24.1 was constructed by digital hybridization method using Keio BAC library [7,8]. These BAC clones were sequenced by shotgun sequencing method. We generated contiguous genomic sequence by combining our data with those generated by other sequencing groups for chromosome 8. For this process, we utilized the BLAST program to identify overlapping BAC clones and confirmed their relationship by dot matrix analysis.

Gene prediction. Gene prediction was performed by using two different programs GENSCAN and MZEF. Predicted exons were analyzed with SMART to identify CUB and sushi domains. All analyses were performed with default parameters.

Table 1
List of partial cDNA fragments for reconstruction of full length CSMD3 cDNA

Fragment	Size (bp)	Exons	Position in coding region[a]	Position in cDNA	Comments
B73	221	1a	5′-158:63	1:221	5′ RACE
B72	279	1a–2	13:291	172:450	
B74	168	1b1–1b2	5′-171[b]:5′-4[b]	1[b]:168[b]	5′ RACE
B71	357	1b2–3	5′-40[b]:437	132[b]:596	
B70	488	2–4	210:697	369:856	
B68	277	4–5	561:837	720:996	
B67	267	5–6	754:1020	913:1179	
B66	880	6–12	933:1812	1092:1971	
B61	138	11–12	1675:1812	1834:1971	
B60	1067	12–18	1811:2877	1970:3036	
B54	397	18–20	2856:3252	3015:3411	
B52	357	20–22	3204:3560	3363:3719	
B51	563	21–24	3463:4025	3622:4184	
B48	237	24–25	3940:4176	4099:4335	
B47	416	25–27	4089:4504	4248:4663	
B45	649	27–30	4396:5044	4555:5203	
B44	797	28–32	4572:5368	4731:5527	
B40	648	32–36	5279:5926	5438:6085	
B36	558	36–41	5825:6382	5984:6541	
B31	476	41–43	6327:6802	6486:6961	
B29	430	43–45	6714:7143	6873:7302	
B27	227	45–47	7066:7292	7225:7451	
B26	242	46–48	7204:7445	7363:7604	
B24	372	48–50	7400:7771	7559:7930	
B22	697	50–53	7711:8407	7870:8566	
B21	661	51–54	7951:8611	8110:8770	
B18	402	54–56	8494:8895	8653:9054	
B16	867	56–61	8855:9721	9014:9880	
B12	449	60–62	9588:10036	9747:10195	
B11	237	61–63	9837:10073	9996:10232	
B10	432	62–65	9956:10387	10115:10546	
B7	1627	65–71	10230:3′-732	10389:12015	
B2	1174	71	3′-435:3′-1608	11718:12891	
B1	401	71	3′-1466:3′-1865	12749:13148	3′ RACE
B0	257	71	3′-1581:3′-1865	12864:13148	3′ RACE

[a] "5′" indicates base position upstream from the initiation codon. " 3′" indicates base position downstream from the last base of stop codon.
[b] Base position based on the numbering of exon 1a transcript (However, −40 and 1 in those columns are counted based on the numbering of exon 1b transcript).

Table 2
Exon/intron structure of human CSMD3 gene

Exon number	Exon size (bp)	cDNA (nt)	Genome position NCBI build 33 chr8 (bp)
1a	336	1–336	114118034–114117699
1b1	81	1–81	114058175–114058095
1b2	148	82–229	114057907–114057760
2	223	337–559	113995815–113995593
3	113	560–672	113959726–113959614
4	195	673–867	113854938–113854744
5	208	868–1075	113779985–113779778
6	113	1076–1188	113700201–113700089
7	312	1189–1500	113657170–113656859
8	78	1501–1578	113635783–113635706
9	88	1579–1666	113628899–113628812
10	125	1667–1791	113602773–113602649
11	122	1792–1913	113540288–113540167
12	104	1914–2017	113510811–113510708
13	113	2018–2130	113481296–113481184
14	183	2131–2313	113371072–113370890
15	327	2314–2640	113366754–113366428
16	195	2641–2835	113363658–113363464
17	139	2836–2974	113347437–113347299
18	188	2975–3162	113337363–113337176
19	189	3163–3351	113331371–113331183
20	117	3352–3468	113326247–113326131
21	216	3469–3684	113319933–113319718
22	189	3685–3873	113318027–113317839
23	170	3874–4043	113268257–113268088
24	157	4044–4200	113254679–113254523
25	192	4201–4392	113237976–113237785
26	127	4393–4519	113233742–113233616
27	203	4520–4722	113231895–113231693
28	192	4723–4914	113198247–113198056
29	139	4915–5053	113187851–113187713
30	188	5054–5241	113184999–113184812
31	195	5242–5436	113173705–113173511
32	117	5437–5553	113153729–113153613
33	114	5554–5667	113090054–113089941
34	96	5668–5763	113089435–113089340
35	204	5764–5967	113087749–113087546
36	125	5968–6092	113071810–113071686
37	97	6093–6189	113064685–113064589
38	105	6190–6294	113061478–113061374
39	119	6295–6413	113033556–113033438
40	70	6414–6483	113032266–113032197
41	117	6484–6600	113027235–113027119
42	210	6601–6810	113022708–113022499
43	189	6811–6999	113018753–113018565
44	178	7000–7177	113017851–113017674
45	146	7178–7323	113016496–113016351
46	81	7324–7404	113001003–113000923
47	114	7405–7518	112999972–112999859
48	189	7519–7707	112995639–112995451
49	147	7708–7854	112995074–112994928
50	189	7855–8043	112992188–112992000
51	186	8044–8229	112987214–112987029
52	195	8230–8424	112985937–112985743
53	174	8425–8598	112982988–112982815
54	174	8599–8772	112977028–112976855
55	174	8773–8946	112973732–112973559
56	186	8947–9132	112972717–112972532
57	174	9133–9306	112970560–112970387
58	183	9307–9489	112968268–112968086
59	177	9490–9666	112962372–112962196
60	180	9667–9846	112946612–112946433
61	174	9847–10,020	112944834–112944661

Table 2 (*continued*)

Exon number	Exon size (bp)	cDNA (nt)	Genome position NCBI build 33 chr8 (bp)
62	174	10,021–10,194	112936449–112936276
63	74	10,195–10,268	112935348–112935275
64	112	10,269–10,380	112928153–112928042
65	180	10,381–10,560	112925595–112925416
66	66	10,561–10,626	112922807–112922742
67	159	10,627–10,785	112918370–112918212
68	113	10,786–10,898	112915499–112915387
69	88	10,899–10,986	112912654–112912567
70	136	10,987–11,122	112909913–112909778
71	1999	11,123–13,119	112905952–112903954

Amplification and sequencing of cDNA between predicted exons. Primers with forward and reverse directions were designed according to the sequence of each predicted exon, which putatively encoded CUB and sushi domains. We attempted to amplify partial cDNAs among juxtaposed exons with those primers in proper combination using cDNAs of adult brain, fetal brain, and testis in Human Multiple Tissue cDNA panels (MTC panels; Clontech) as template. We then examined possible connection among these putative genes by PCR. The final sets of PCR primers to recover the entire cDNA of CSMD3 gene are listed in Table 4. Amplified cDNAs were sequenced with appropriate primers and BigDye Terminator Cycle Sequencing Reaction Kit (Applied Biosystems, Foster City, CA) using ABI Prism 3700 DNA analyzer. Quality of the sequence was proven with Phred/Phrap/Consed program. Homology plot analysis was performed using DNASIS-Mac software version 3.5 (Hitachi Software Engineering, Japan).

RACE PCR. The 5′ and 3′ sequences of the CSMD3 gene were determined by RACE using Marathon Ready cDNA kits of human fetal brain or adult testis (Clontech) according to the manufacturer's instructions. For exon 1a, first 5′-RACE amplification was performed between adapter primer 1 (AP1: CCATCCTAATACGACTCACTA TAGGGC) and 5′ anti-sense primer (CTTGCCAGGCTCCCAGGG TTTGGATTCCTT) using human fetal brain RACE template, and then second amplification was performed between adapter primer 2 (AP2: ACTCACTATAGGGCTCGAGCGGC) and second 5′ anti-sense primer (GCAGTGGAGTTGTTGCTGTTGTTGGTG). For exon 1b, first 5′-RACE was performed between AP1 and first 5′anti-sense primer (TGCTATTATTACCCATGTGCAG) using human testis RACE template, and then second amplification was performed between AP2 and second 5′ anti-sense primer (ACCGGGGTTGTG CTGAAGAGGTTGTTTA: this primer was designed according to a draft version of AC055788. Later, this original sequence was updated to be ACCGGGTTGTGCTGAAGAGGTTGTTTA in the finished version of AC055788. However, the original primer worked to obtain proper RACE product). 3′-RACE amplifications were performed between AP1 and 3′ specific first sense primer (CTGCTTGGCTCTATC TCTACAACATTTA), and then performed between AP2 and 3′ specific second sense primer (TTTTAAATTTTATGCCATCAGTTATT). The PCR products were purified and sequenced.

Table 3
SNP candidates in CSMD3 gene region

Amino acid change	Exon/intron	Position of nucleotides	Allele frequency	NCBI SNP cluster ID
None	1a	-78 G/A	3/11	None
None	1b2	IVS1b2+22 T/C	6/8	None
None	1b2	IVS1b2+216 A/G	1/13	rs4876512
I219M	4	657 A/G[a]	30/20	rs2219898
None	4	IVS4-58 A/C	1/13	rs2030506
None	4	IVS4-5 T/A	13/1	None
None	11	IVS11-21 C/T	10/4	None
None	14	IVS14+56 A/G	1/13	None
None	15	IVS15+68 C/T	6/8	None
None	18	IVS18-5 T/G	4/10	None
None	19	IVS19+91 C/A	4/10	None
None	23	IVS23+88 C/T	7/7	None
None	23	IVS23+99 G/A	7/7	None
None	27	IVS27-99 A/T	3/11	None
None	28	IVS28+30 C/T	12/2	rs2853244
None	29	IVS29-3 C/T	2/12	None
None	31	IVS31-78 T/G	12/2	None
Y2068Y	39	6204 T/C[a]	33/17	None
None	39	IVS39-38 A/G	7/7	None
None	45	IVS45-208 A/G	9/5	None
None	53	IVS53-8 T/C	7/7	rs4876462
None	56	IVS56-169 A/T	7/7	None
None	57	IVS57+46 T/C	7/7	rs1861755
None	61	IVS61-115 A/G	7/7	rs1861753
None	61	IVS61-53 T/C	2/12	None
N3621H	70	10861 A/C[a]	5/9	rs1592624
None	70	IVS70+21 C/T	6/8	None

[a] Numbers from the first base of initiation codon of exon 1a transcript.

Mutation search. PCR primers of 50–250 bp flanking exon–intron boundaries were designed manually or by the program Primer3. Genomic DNA (10 ng) from patients or control individuals was used for amplification of each exon. For exons of more than 400 bp, two or more pairs of primers were designed to keep size of PCR products less than 500 bp. All primer sequences are listed in Table 5. The PCR products were sequenced by ABI Prism 3700 DNA sequencer. These sequences were assembled and searched for SNPs by Consed with polyPhred [9].

Multiple-tissue cDNA panel and Northern blot analysis. Expression analysis of CSMD3 gene by PCR was performed using MTC panels (Clontech; I, II, fetal and immune system panels) consisting of cDNAs from 27 human adult and fetal tissues with primer pairs of A:(CGC AAAGGGGAAAGCCGAGCAAAGGA/CTTGTCACTGGAGGT GGCAGATGGAATC), B:(TTGCGACTCATAAACAACCTCTTC AG/CTTGTCACTGGAGGTGGCAGATGGAATC), and C:(AGCT CGCCTAATGTTACGCATA/AAAGCTGCTTGGCCATTGTTAT) as shown in Fig. 1. Twenty picograms of cDNAs was used as templates.

Northern blot hybridization was performed with Human MTN Blots 1-3 (Clontech) loaded with 2 μg of poly(A)$^+$ RNA from 27 human adult and fetal tissues. Twenty nanograms of partial cDNA fragment amplified by PCR with a pair of primers (AGGACTTAA TGGCACTATAGAA/GAACAGGAAAATCCCACGAAGC) was labeled with [α-^{32}P]dCTP by random primer extension method. Hybridization was carried out according to the manufacturer's protocol. Autoradiography was carried out using IP plate (Fujifilm, Japan) for 24 h and then hybridization signals were detected by FLA-3000G system (Fujifilm, Japan).

Sequence analysis. Homology search of nucleotide and amino acid sequences in the database was performed through the BLAST server at NCBI using BLASTN, BLASTP, and BLASTX. The fugu genomic DNA sequences corresponding to CSMD gene family (Scaffold_3752; CSMD1, Scaffold_380; CSMD3a, Scaffold_55; CSMD3b, and Scaffold_73; CSMD2) were downloaded from Ensembl genome server and analyzed by DOTTER program to compare with cDNA sequences of human CSMD. Identified exons were joined to construct putative cDNA sequences and used for further analysis.

Evolutionary analysis. The amino acid sequences deduced from the coding sequences of CSMD family genes of human, mouse, and fugu were aligned using the Clustal X (ver 1.82). An evolutionary neighbor joining tree was constructed by Clustal X. The phylogenetic tree was drawn by TreeView (ver 1.5). Graphic representation of the aligned amino acid sequence was obtained by BOXSHADE 3.21.

GenBank accession numbers. Accession numbers of BAC clones covering the entire CSMD3 gene are as follows: AC055788, AC107890, AP005015, AP004713, AC027189, AC021264, AP002846, AP003470,

Table 4
Primers used for amplification of CSMD3 cDNAs

Exon	Forward primer (5′–3′)	Reverse primer (5′–3′)
1a	AP2[a]	GCAGTGGAGTTGTTGCTGTTGTTGGTG
1a	AP1[a]	CTTGCCAGGCTCCCAGGGTTTGGATTCCTT
1a–2	CGCAAAGGGGAAAGCCGAGCAAAGGA	TGCTATTATTACCCATGTGCAG
1b1–1b2	AP2[a]	ACCGGGGTTGTGCTGAAGAGGTTGTTTA
1b1–2	AP1[a]	TGCTATTATTACCCATGTGCAG
1b2–3	TTGCGACTCATAAACAACCTCTTCAG	CTTGTCACTGGAGGTGGCAGATGGAATC
2–4	AGGACTTAATGGCACTATAGAA	GAACAGGAAAATCCCACGAAGC
4–5	TGTATTATATGGCACAAGATTC	AATTGTGTCCCCAGGCTCTGCT
5–6	ATATCCAGCCCTAGTTTTCCTA	AGCACTAAATCCACGGTATCGA
6–12	TATACCACCACCAATTATCAGC	TGTCAAGGTATCATAGCCAATC
11–12	TTTACATCTCCCAACTTTCC	TGTCAAGGTATCATAGCCAATC
12–18	CAATTGGCGATGGGGGCGAAGT	TGACAGAAGATTTGGCCCATCA
18–20	TGATGGGCCAAATCTTCTGTC	CGGGTAACCAGGTGATAAGATT
20–22	AGGAGATGTTAGAGGGCCTAGT	ATGCCAGGATCTTCACAAGGTT
21–24	GCTCAATTGCGTTTCATTTCA	AAGCCTTCATCTGTCCCTTCAG
24–25	TCTATGCGCGGACTGACACTTAGT	AAGGCTACTTCCGTGGAGAGTG
25–27	ATACAAGATCAGTGACCAAGGC	TTACTATATTGAGGGTGCTATG
27–30	CATGATATACTCCGAGTCTGGG	CATCACTCCGAAAAGCCAAATG
28–32	TGCCACTGCGTGTCGTGACCCA	AAACACAATTATGTCCCACACT
32–36	CGCCCTGTGGAAGTCGTTCAAC	CAGCTCCCTCTGGCACTGTGAT
36–41	TTTTAACTAAGCGCAAAGGGAC	GAAACCCAGGACTGAGGATCAC
41–43	TCAGTGTGGTGGTGCTATGTCA	CTCCGTGAAGGCATGTGAGAGC
43–45	TGATTTTACTGTGGGTCAAACCA	GCCACTTGTTGTGAAATCACTG
45–47	AATACCGCTTTGGAATCAGTCT	CCAACTAAAGTAAATCCTGGAA
46–48	CCCAATGCTGAAATTTTGACGGAAGATGAT	TGAAGATTTGGGTAACTGTCAG
48–50	GAGTCATATTGAGCCCTGGAT	CACTGTTAAGCTGCCCACCTGT
50–53	ACACCAGAATCCCCACCTCATG	CTGAGGAAAGGCATTCCCTTAC
51–54	GGAACGCGAGTTACCTATTTTT	CACAGGATGGGAGCTGACCAGA
54–56	AATTATGGATATAGAGACACAG	AGGATTGCAGTCATAGAATAC
56–61	ATTTTACTTACGGCACTGTGGT	TAGAGATCTGGGGAGGAGTTG
60–62	TTACATGTGCCAGCCAGGCTAC	CTATGCAGTGAGGTGATGAACCACTCCAAG
61–63	GAGTGGTGAAGTACCGCAGTGC	GGCACACCTGGGTTTTCACAAG
62–65	TTCCTTTCATATTAGTGGGATC	CTTTTCCAGTCCAGGTGTTATC
65–71	CTGTAAACAGCCAGAAACTCCT	TACCCCATAGCTCACTGAATAA
71	ATCACGCCCACAAATCTGTCTT	TAAATAACTGATGGCATAAAAT
71	CTGCTTGGCTCTATCTCTACAACATTTA	AP1[a]
71	TTTTAAATTTTATGCCATCAGTTATT	AP2[a]

[a] The sequences of AP1 and AP2 which were used for 5′- and 3′-RACE are described in Materials and methods.

Table 5
Primers used for mutation search

Exon	Forward primer (5′-3′)	Reverse primer (5′-3′)	Size (bp)
1a	CCGATTCATTATCGTCACGGTTT	GAACGAGCTGTGAATCAACTCCTTTAGTAT	803
1b2	CGCAGCGCACACACCTTTAGTTT	GGGTTGGAATGTACAAGACACTTAAGAGTA	502
2	AATTATTGTTAACTTTGGTCTTGGAAATGC	GGTGAATACAACTTTATTATAATTGCATCA	450
3	TGTTGGCTGGATAGCATTCCTTCTGGAAAT	GAGACAACATGATGTTGTCTCAAATGTTGA	358
4	TGACTCTTCTCCAAAGAAATAAGAATATCC	CTCTATCCTTTAGATTCCGTAGAAGAAACA	411
5	TATTATGCACTTGGTATCACATTTGAACTG	CAACCTCTTCACATAACTTTCCTGTAAGAT	468
6	ATGAATATAGTTGGACATTTGTTTGGTCCC	CTCAGCAACAGCATGACCATTTTTTTCAAT	351
7	TCTTCATGTTGATTATCCTTAAGGTAATAG	GGCTCTCTTTCAGCTCATTCTCTAATTAGA	523
8	TAAGAAAGTGGTTTCTTTATCTTGCATAAG	GCATGTTACACAGATATTAATGCAGTTAAT	325
9	TAACAAATGCTTAAATGACAGTGATTATAG	GTTTGATGCCACTAAGACTTCATCTTTATA	331
10	TACATATTTGACCTATTGGATATTGCCAGC	GCACAATCTTTCCTTAAAGATTATATTACA	340
11	AATGTTGAAAATATCGAGCTTGTGAATCCC	CCCAAATTTCTCATATAAATCTGGTTTATA	491
12	TATTCTAACATGACTTGCTGCCATACTATG	GGAGCGATCAATGTAAAACAATAATTTGAA	391
13	TCAACACATACTCCTCTATGTTTATCTGAC	GGAGGAAATGATCTTTGTTGAAATGTAGCA	386
14	TATCTCCACAACAAAGAGCATATATCTACC	GCAGCTCTTTGCAATTAATTACAAAAGCAA	368
15	TAGAGCCTTAAGTTTATCCAAAGTACTAGT	GGCACAATGAGATTTACAGCTTTTACAAAT	660
16	TCTCTCTTTTCCAAAATATAGAAGGATCTC	CTTGAACCTTGTTGACTATATTTAATTGGA	796
17	ACAGAAGTAATTGACTAGATGTACCATTGC	GGAAGCATTCAAAGCACAGCAATTGACTAT	428
18	AATTTAAATAGCAGTCCAAAGCATTATTTC	GCAGAATTTACTAAAACACACACCTTATAA	437
19	TCATTGCTTGTAATTGTTTGAAGAACTAGG	GGTTTGATAACATAAATTTGCCAGAAATCA	429
20	AGATAATGCGACTGCTCAATTCAAGAGAGG	GCTCATGCAAAATAACATGTAAAGGAAAGT	312
21	TATTTGGAATATTTTATAGCAATGCCTCTC	CCAGCTATTTTTCACTGTTTTATCATCAGT	460
22	TACGTTCTGATAGCTATTCATATAGTCCCT	GCAGAAATTCAACAGATTATAAAGAACGAA	432
23	AATGCATAACACTTCCATACATGATATTGC	GTGAATGGTAGAATCAACTTATGGTTAATA	389
24	TTAGTTGCCCTTTGGAAATGTACTGATCTG	GGATATCATTTCAATTTGTGCAATGTAATT	378
25	TATGATACGAAAATTAAATATGGTAAGGGC	CAAGCTATGAGTTCATTAATTTCTCGATAT	472
26	TTCTTTTGTCTACTCATGAGAAATTTGTAG	GGGTAATAATTTGCAAAACTCTGTATATCA	635
27	ATAGCAAAAAAGAACTATCTAGTAAAAGGC	CCTTTGCTGCAGAAATATAGGAAGACTAGA	460
28	ATACTATGCCACTTTTTCAAATAAGTTAGC	GACTCTCATATAGAATCTAGTCTAGAATAT	482
29	GCCTCATAAGCTCTTTTAGAGACGTGAT	AATGCTAAAGCAAAATGGCCTAACAATATA	304
30	GTTGCCAGGTGATTTTTTATATTAGCTTAT	ATACTAATTGTTAATTTGTTTGCCTGATAC	591
31	TTTTAAAGCAATTAATGTTATACTAACACT	TAGTACGACACTTGTAGATTGTATGTGGAA	353
32	AATAAATTATACACCTTAAATTCCATTAGC	ATGTAGTGAAAAATAAAGCATGTCTG	297
33	TTATGACCTGAAGTACATGTGTCAGTAA	TTCAGTTTGCTTTATTTCTTATATTGATGA	417
34	TGCAATACAAGTTGAACCCAAAGGA	GAGGGGACAAAATAAAACTGCTTAAAAACA	397
35	AATGTATTCAAATTGTTAATTCAAGTTCTG	AAATCCATATAAATCAGTAAATGAGGAAAT	585
36	TGTCTAAGATGTTATCTCTGACCTTTACTC	CACTTGCTAAACTTTTAGAAATAATGTTCA	310
37	GGATTTGAACTCTCTTTTCCAAGGGGGATT	AATTAAAACTACCAAGTAAGAGTTGTAGAT	291
38	AACTTGAATAATGCATATTTTTAGTGGTAT	TGTAATTTTGCAATTGAACTGAAGTAATAG	385
39	GAGACACTGATATTCAATTTTAGCAACTTA	CAATGTTGTAAAACCCGTGGCTTATCAATC	309
40	AATATGTGTATAAAAATAAAGCCATGTACT	ATTTCCATAAGATCTGGTAAAAGGCAAACT	347
41	ATTCCTAGTGTATAGTGAGTGAACAGACAA	AAGGCTACATTAATAAATATTGAAAGCTCT	323
42	TAAATGCAAACTTATCTAAGTACATTGATG	CTAAAAATCTTCCTAATTGAAGCTCACCTT	466
43	AAGTTGCAAAATTGCTGTAGTACCCTATCT	ATTGGAAGAGGAAGACAAGTAAAAAGTGCC	360
44	ACAAGGCATAAATGTGTTTCATAATGCTTA	TTTATTCCAACCTTGTTTTACTGTGTATAT	358
45	AATATATGCAATACTTTTATTAAATACGGT	TTTCCAGGACAAATTAAACAGTAGCAAATG	293
46	TTTTGAACCACATAACAAGAAGGACTAGAA	GAATGAGGGTCTCAAGACAGGCTAT	443
47	CTGATGGGTCAATGAGTATTATATACTG	TTATAAAATGTATAGCATTGTTTCCAGTTA	435
48	GAACATGAAAGCAGAATACTCAGTATCTAA	GTATAATTTGTAAAATAATAGTGGAGCTAA	474
49	AGCAATTTAACCATATTTAATACTATTCTA	AAATGCTTTTAGGACATATAGGATAATTTA	524
50	AGCACTAGTCGCTGTATAGGTGTTTATCAC	AAAACGTTAAATGGTGCTAATCTCACATTC	341
51	TCAGAGGAATTGATAAAAATCCCCGTTAGT	AGACATCACATTTCAGTTTAGGGATTGGTT	418
52	TCAGTGGAGATGTAAAGAATGGAATAGACG	GATTCAAACAATTCGATAACCAAACATAGC	337
53	TGTAAATACATCAAATCATAGGAAGCTCAA	AGTATTTTCTGGTTTTTAACATACTTACTG	414
54	CATATAGTTATAATTGAATTATTTGTTGCT	ATAACTGAAATTCTCTTGAACATATTCCAC	497
55	CGTATATTTCCAGATAAAAAGTAGTCTTCT	TCTGCTCAAGCTGTTTTAAGTGTGTCACTG	397
56	TCTTCTGTATTTTCTACCACATAGATGTCA	ACAATAAAATCAATATGCTTCTCTGTGATA	496
57	ATCCTCGGAATGAGTAAGTTAACAGAACAG	ATTGTTTTGTGTATACGTTGCTACTACTAC	470
58	ATGAATTACAATTTTACTCCATCTGTGCTA	GTAGTACTCATCTGGATTTAGATACACACT	393
59	TGAATAATTTGAAAGAATCGTTTAGTCATT	GCCATGGTAACTACTATAATCCTTTAATCT	427
60	ATTAAGTATGCCATATGCTTACTACTCCTC	AGTACTAGAAAAATAAATGCAACCTGGAAC	392
61	ATGTCACAACATATTGTCCTCGAAAAGGTT	AGGCATTGAAGGAAGCTTAACAATCTCATA	391
62	ATTCAGTATTTATCTGCACATAAGCTAAAA	AAATGTCACCCCTATTAATTACACAGTTAC	409

Table 5 (continued)

Exon	Forward primer (5′-3′)	Reverse primer (5′-3′)	Size (bp)
63	TTGCTTCAAGATTAACAGGCTTCTGTCCAT	GAACCAAAGACGCATTCTGACCTAGTATGA	307
64	GATAACATTTTCTCACTCAAGCTCTGTTTG	AAAATGCATATAAATTCTTACCCATGATAG	254
65	TATATTTGCATTGAAACATTTTTGTGATTT	AGGAAAGCAATAGTGCAATCTCTTGTGTTA	397
66	GCATCAACTTTTGTAAACACCTTTATTTCT	ACTATGAACAATACTTTTGGAATACTAACA	306
67	TAAACCTAATATCTTTAACTGGTGTTTTCT	AAACATTTCACCATATAAAAAGCATTACTA	511
68	TAATTTAAACCCCACAACTAAGAAATAAGT	GACAAAAGTAAGCAAAACACATCAGAATCC	543
69	TAAAATTGTGTATTTATTAAGGCTGGGAAA	TTTATAATGCAGTAAGTGTACCTAATGTAC	273
70	TATTTTTGTTTACTATGTGATTTCCAGGTA	TATGTTGTATGTAATGTTGCAGGGGTGGTT	452
71	TATTTTTAAGTTACATTATCTGCTAACGAT	AAAGGTGACCCAGCAAGTCCTGAAAAGTGT	382

AP005017, AP002979, AC087710, AC024996, and AC007719. Accession numbers of cDNA sequences of the CSMD3 gene are AB114604 (CSMD3a) and AB114605 (CSMD3b).

Results

Identification of CSMD3 gene

We have sequenced and analyzed a BAC contig of 30 Mb corresponding to the 8q22–q24.1, within which the genetic locus of an epilepsy BAFME1/FAME has been mapped (Figs. 1A and B). Extensive computer-aided analysis with exon prediction programs revealed many potential exons, some of which encoded CUB and/or sushi domains (Fig. 1B). We integrated these predicted exons and formulated initially eight putative genes (CUB1–CUB8, data not shown). CUB1 showed significant homology to a cDNA clone KIAA1894, but no other putative genes had any significant homology to the cDNA sequences deposited in GenBank nr database. We then designed a series of PCR primer pairs to amplify every exon of the eight putative CUB-containing genes and attempted PCR amplification using human cDNAs of adult brain, fetal brain, and testis. Moreover, we examined possible connection among these putative CUB-containing genes by PCR and found that all the neighboring two genes were actually connected to each other to form a single gene. Consequently, all the partial cDNA fragments were integrated into a single contiguous cDNA sequence of 13 kb (Fig. 1B, Tables 1 and 2). At this stage of analysis, we realized that this is a giant gene homologous to CSMD1 gene consisting of CUB and sushi multiple domains which was on human chromosome 8p23 [10] and hence we tentatively named it CSMD1L gene. However, as described later, we identified another CSMD1-like gene on human chromosome 1p34 and this new gene was more closely related to CSMD1 gene. We therefore designated this gene as CSMD2 and renamed CSMD1L as CSMD3 gene.

PCR amplification analysis using various tissue cDNAs indicated that there are two alternative first exons, namely exons 1a and 1b. In fact, two remarkable CpG islands were found in the genomic sequence of corresponding region (Fig. 1C). Transcripts in which exon 1a was connected directly with exon 2 were found in adult and fetal brains, whereas transcripts in which exon 1b was connected with exon 2 were found in testis. However, no cDNA fragments connecting exons 1a and 1b were amplified from any tissues in the MTC panel, and hence we concluded that exons 1a and 1b act as an independent first exon, generating two distinct transcripts. The first base of exon 1a was determined by sequence analysis of the longest 5′-RACE product from fetal brain cDNAs. We also analyzed 5′ end of exon 1b using testis cDNAs and unexpectedly we found an additional exon very close to exon 1b. Therefore, we named the original exon 1b as "exon 1b2" and the new exon as "exon 1b1", respectively. Thus, this new gene consists of 73 exons and produces two transcripts from two independent first exons (1a and 1b1) (Fig. 2A). We also found in the transcript of brain that exons 32 and 35 were connected and therefore two exons 33 and 34 were skipped, generating a shorter transcript. In addition, we found that exon 7 was spliced out in the transcripts of testis. In adult and fetal brains, two transcripts with or without exon 7 were observed. The longest coding sequences are estimated to be 11,124 bp for the transcript with exon 1a and 11,004 bp for the transcript with exon 1b. All the exon–intron junctions follow the GT–AG rule.

Protein structure of CSMD3

The longest open reading frame of 11,124 bp in this new gene CSMD3 encodes a protein of 406 kDa consisting of 3707 amino acids (Fig. 1F). Examination with SMART program revealed multiple units of CUB and sushi domains. The protein deduced from a transcript variant without exons 33 and 34 lost 10th CUB domain and therefore the predicted molecular mass was decreased to 398 kDa.

There are 14 CUB domains which are located from N-terminus and intervened by 13 sushi domains (Fig. 1F). There are additional 15 sushi domains which are arranged in tandem starting from the last CUB domain toward the C-terminus. A transmembrane domain was predicted at a position after the last sushi domain. The transcript which started at exon 1b caused

no change in domain structure (Fig. 1D). These analyses indicate that CSMD3 protein is a transmembrane protein composed of a long N-terminal polypeptide with CUB and sushi multiple domains exposing outside the cell, a single transmembrane domain, and C-terminal domain of 55 amino acid residues located in the cytoplasm.

Expression of CSMD3

Northern blot hybridization was performed using human adult and fetal multiple-tissue blots with 5′-partial cDNA of CSMD3 (containing exons 2–4) as a probe (Fig. 2A). The Northern blot analysis indicated that the longest transcript of about 13 kb is detected mainly in the adult brain, fetal brain, and testis (Fig. 2B). Some shorter transcripts were also found in the testis and fetal brain. The 13-kb transcript was not clearly detected in other tissues, but smears were seen in some tissues such as pancreas and spleen.

We then made further expression analysis by PCR amplification of cDNA fragments in the MTC panel. For this, we used three sets of PCR primers, which are designed to detect any cDNA fragments with the sequences covering exons 1a–3 (primers A), exons 1b2–3 (primers B), and exons 68–71 (primers C) (Fig. 2C). The

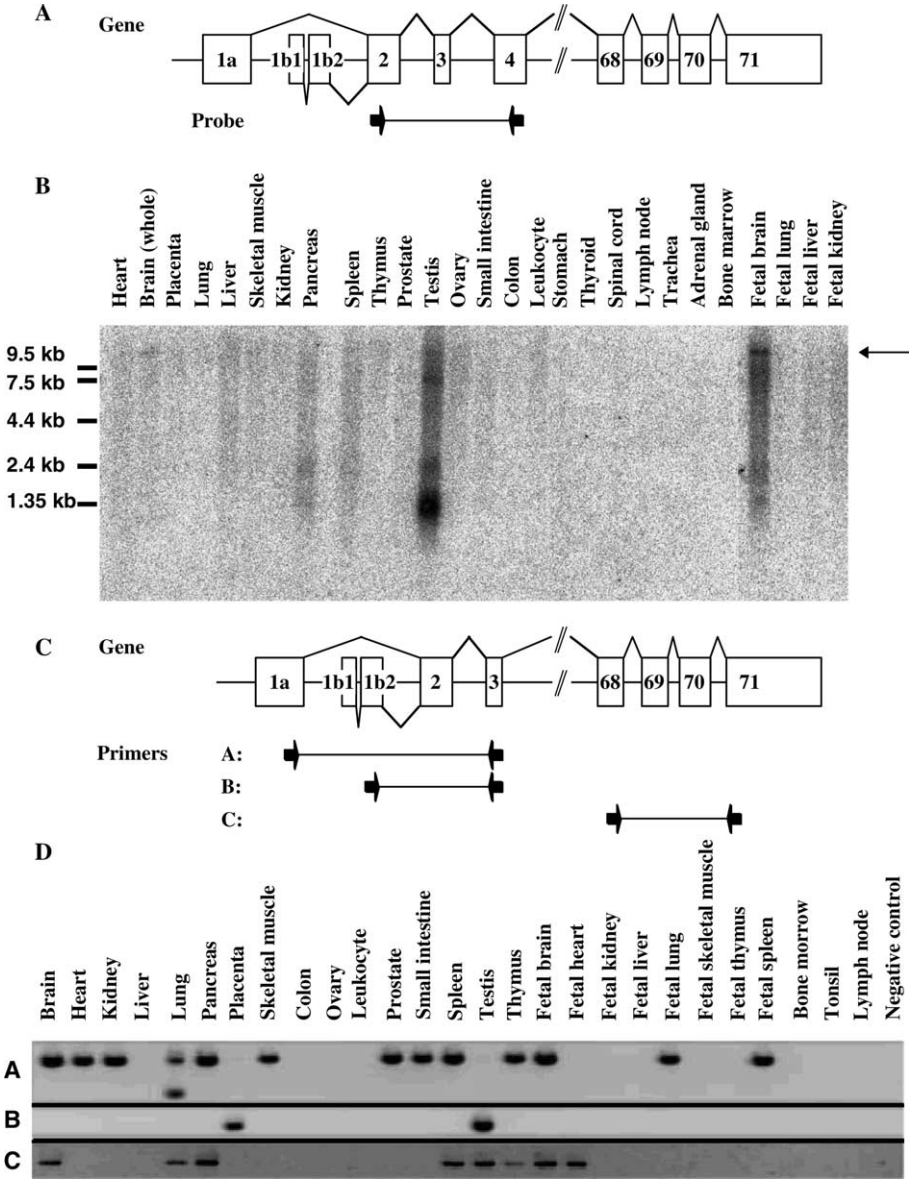

Fig. 2. Expression of human CSMD3 gene in various tissues. (A) The position of a cDNA probe used for Northern blot hybridization. It detects exons 2–4 of CSMD3 gene corresponding to the N-terminal CUB and sushi domains. (B) Northern blot analysis of CSMD3 gene expression using human multiple-tissue Northern blots of 27 tissues. Arrow indicates the longest transcript of CSMD3 gene. (C) Three sets of PCR primers (A, B, and C) used for expression analysis of CSMD3 gene. (D) PCR products amplified from each tissue in the MTC panels with three sets of primers. A second band in the lung with primer set A may represent an alternatively spliced transcript variant.

cDNA fragments covering exons 1a–3 were detected in many tissues including adult and fetal brains, whereas those covering exons 1b2–3 were detected only in placenta and testis (Fig. 2D). Interestingly, cDNA fragments covering exons 1a–3 were not found in placenta and testis, indicating the alternative use of first exons 1a and 1b. In seven tissues including adult brain, fetal brain, and testis, cDNAs covering from 5′-end through 3′-end were detected. These results are not perfect match with the Northern blot data, but strongly support that CSMD3 gene is expressed mainly in adult and fetal brains. The expression in the adult brain led us to examine CSMD3 gene as a candidate gene for the BAFME1/FAME.

Mutation analysis

We investigated the possible sequence variations in the coding regions of CSMD3 gene by PCR-based sequence analysis. Based on the exon/intron structure of CSMD3 gene, we designed sets of PCR primers to amplify every 72 exons except 1b1. For some larger exons, multiple overlapping PCR primers were generated. Using these PCR primers, we analyzed seven patients from five families

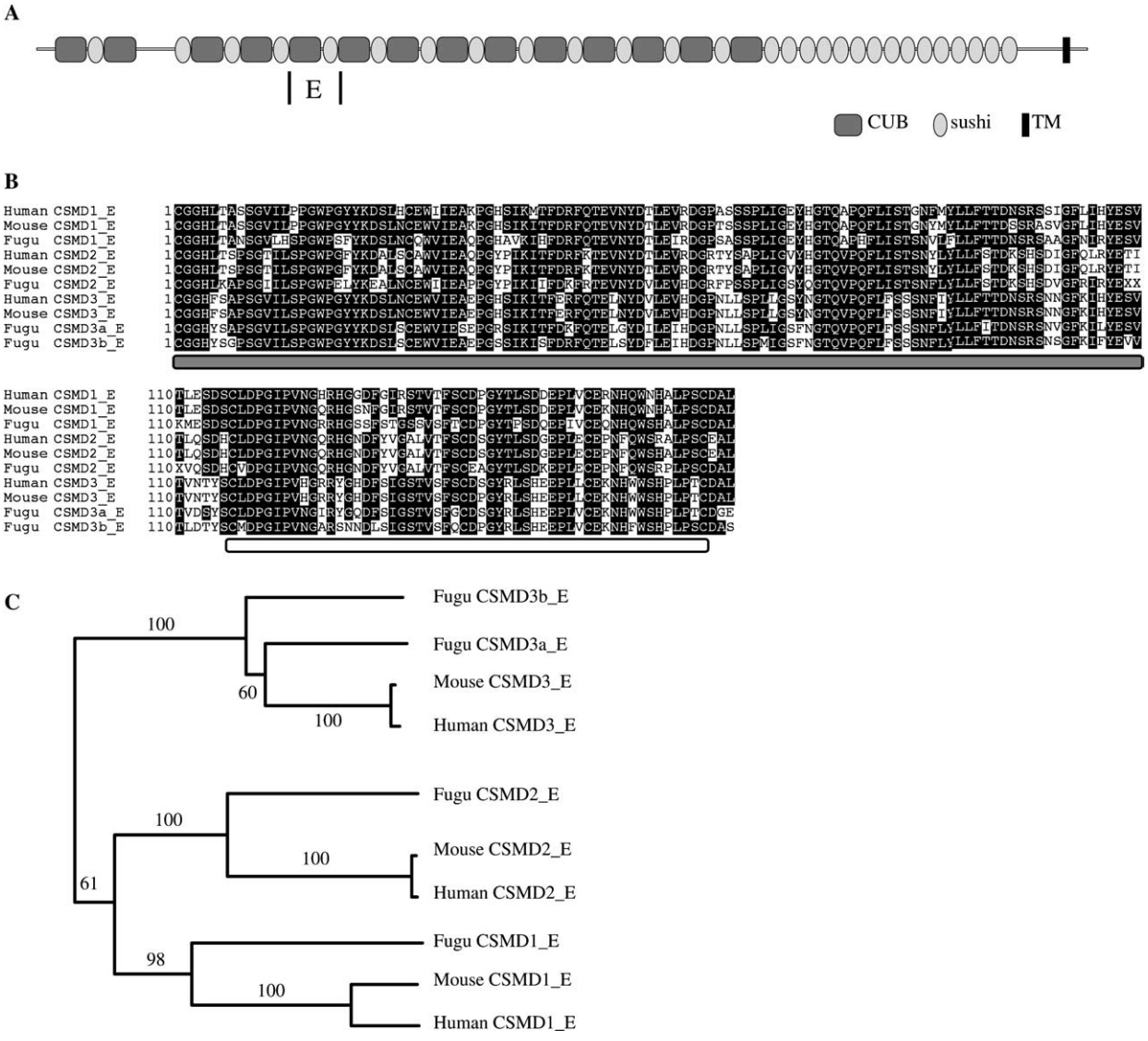

Fig. 3. Comparison of CSMD3 gene families among human, mouse, and fugu. (A) The position of the 5th unit of CUB and sushi domain (E unit) in the CSMD3. (B) Alignment of the E units of CUB-sushi domain in the CSMD gene families of human, mouse, and fugu. Alignment was preformed by using the BCM implementation of Clustal X. The CUB and sushi domains are underlined with thick gray and white bars, respectively. Black background indicates that these amino acid residues are identical for more than half of the aligned amino acid residues. Three "X" letters at amino acid residues 108–110 in fugu CSMD2_E indicate undetermined amino acid sequence remaining in the fugu genome scaffold sequences. (C) Evolutionary trees of CUB-sushi unit E among human, mouse, and fugu CSMD. The numbers labeled with the bootstrap nodes indicate percentage of reliability of this tree deduced from 1000 bootstrap replicates.

with BAFME1/FAME and identified three types of single nucleotide changes, such as 6204T/C, 10861A/C, and 657A/G in the coding sequences (Table 3).

The 6204T/C transition was found at the third letter of Tyr codon, but caused no amino acid substitution. This single nucleotide polymorphism (SNP) was a new type and was not deposited in the NCBI SNP database. The 10861A/C transversion caused amino acid substitution from Asn to His (N3621H). Interestingly, two patients from two unrelated families showed homozygous change of this type 10861A/C, however such a homozygous change is unlikely in the case of autosomal dominant disease like epilepsy BAFME1/FAME. The 657A/G transition caused amino acid change from Ile to Met (I219M). We expanded the mutation analysis of 6204T/C and 657A/G for additional 18 members of one of the patient families. However, these changes were not associated with the patients in the family and in fact these substitutions, 10861A/C and 657A/G, were found in the SNP database. Thus, we conclude, these three types of base-substitution are not responsible for pathogenesis of the BAFME1/FAME.

During the course of mutation search, we identified 24 potential SNPs including 18 new SNPs for flanking sequences of 72 exons of CSMD3 gene. These SNPs should be useful for further genetic analysis, particularly for linkage disequilibrium mapping of the region.

Comparison of CSMD genes among human, mouse, and fugu

We performed a homology search to the genomic sequence databases NCBI and Ensembl using the sequence of CSMD1 [10] and CSMD3. We identified an additional CSMD gene (CSMD2) on human chromosome 1p34. We also found three CSMD homologs in the mouse genome and four CSMD genes in the fugu genome. Thus, we could construct putative structure of human CSMD2, mouse CSMD2, and 3. Fig. 1F shows protein structures of human CSMD protein family. The numbers and positions of CUB and sushi domains and transmembrane domain are completely conserved among the human CSMD family. Fig. 1E shows comparison of exon structures and corresponding protein domains among human CSMD family genes. The exon structures are also well conserved except for exon 7 and last exon of CSMD2 and exon 7 and 8 of CSMD3. The exons 7 and 8 correspond to a portion of CSMD3 protein between second CUB and sushi domains. Consequently, CSMD3 protein is longer than CSMD1 and CSMD2 proteins. Interestingly, exon 7 of CSMD3 gene was subject to alternative splicing. We could not construct the entire structure of fugu CSMD because information on fugu genome sequence is incomplete. Nevertheless, using the sequences of human and mouse CSMD gene family and all available fugu sequences, we could identify the fifth unit (designated E unit) of CUB and sushi domain in the CSMD genes of human, mouse, and fugu (Fig. 3). Using the deduced amino acid sequences of the E unit of CUB and sushi domain, we were able to draw a phylogenetic tree of CSMD gene family. A relationship between CSMD1 and CSMD2 was found to be much closer than CDMD3. In addition, we found a fourth CSMD gene in fugu genome which could be derived from fugu CSMD3 gene.

Discussion

We identified a novel giant gene encoding a transmembrane protein with CUB and sushi multiple domains on the human chromosome 8q23.3–q24.1 in which benign adult familial myoclonic epilepsy type 1 (BAFME1/FAME, OMIM:601068) has been mapped. This giant gene consists of 73 exons and spans over 1.2 Mb on the genomic DNA region. Deduced amino acid sequence of the protein showed high homology to two other CSMD genes (CSMD1 on 8p23 and CSMD2 on 1p34) and hence this new gene was named as CSMD3. The CSMD3 gene was expressed mainly in adult and fetal brains as several forms of transcript variants. Comparative genomic analysis revealed the conserved family of CSMD genes in the mouse and fugu genomes.

There is little information about the function of CUB-containing proteins, although it has been postulated that they would be mainly involved in developmental process [11]. On the contrary, some sushi repeats are known as CCP domain, which exists in a wide variety of complement and adhesion proteins and is known to form a β-sandwich arrangement [12].

Epileptic seizures are induced by abnormal electrical discharges in the brain. Recent studies have revealed that membrane proteins for potassium, sodium, and calcium ion channels are the "gates" to regulate neuron signaling and are considered to be involved in some types of epilepsy. For example, it is reported that mutation of potassium gate proteins KCNQ2 and KCNQ3 causes benign familial neonatal convulsion [13,14]. Since ion channels are fundamental in generation of membrane potential, mutation of genes encoding some types of ion channel proteins may cause epilepsy. However, mutation of non-gate proteins also causes epilepsy or seizures in some cases. For instance, cystatin B, which is a widely expressed cysteine protease inhibitor, is responsible for a severe neurological disorder known as progressive myoclonus epilepsy of the Unverricht–Lundborg disease (EPM1) [15]. More recent study shows that SEZ-6 containing both CUB and sushi domains is involved in signaling and cell-cell adhesion. The SEZ-6 was first identified as a seizure related gene by differential screening of mRNA from cortical neurons

treated with pentylenetetrazole (PTZ), which is a drug to induce epileptic seizures. The SEZ-6 is strongly expressed in the developing forebrain and necessary for the formation of neural network [16]. These findings led us to examine CSMD3 as a candidate gene.

Although our initial trial to identify causative mutations in the CSMD3 gene was unsuccessful, a possibility of CSMD3 gene as a candidate gene for BAFME1/FAME still remains because many other types of mutations have not been analyzed due to unusually large size of CSMD3 gene. As a similar example, majority of mutations in the Parkin gene, which is one of the major familial Parkinson's disease genes, are exonic deletions, and such exonic deletions are often difficult to find, particularly in the case of patients with compound heterozygosity, when simple PCR-based method was applied [17]. The size of CSMD3 gene is as large as Parkin and dystrophin, therefore it is likely that exonic deletions might frequently occur in the CSMD3 gene as observed in those giant genes. Because BAFME1/FAME is an autosomal dominant inheritance, and judging from the pedigrees of patients, it is very unlikely that both alleles harbor causative mutations.

Further genetic linkage analysis with newly identified SNPs will be necessary to identify the causative mutations for BAFME1/FAME.

Acknowledgments

We thank T. Asakawa for her excellent technical assistance. This work was supported by the Fund for "Research for the Future" Program from the Japan Society for the Promotion of Science (JSPS) and the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT); and a Grant-in-Aid for Scientific Research from JSPS.

References

[1] A. Kuwano, F. Takakubo, Y. Morimoto, E. Uyama, M. Uchino, M. Ando, T. Yasuda, A. Terao, T. Hayama, R. Kobayashi, I. Kondo, Benign adult familial myoclonus epilepsy (BAFME): an autosomal dominant form not linked to the dentatorubral pallidoluysian atrophy (DRPLA) gene, J. Med. Genet. 33 (1996) 80–81.

[2] N.M. Plaster, E. Uyama, M. Uchino, T. Ikeda, K.M. Flanigan, I. Kondo, L.J. Ptacek, Genetic localization of the familial adult myoclonic epilepsy (FAME) gene to chromosome 8q24, Neurology 53 (1999) 1180–1183.

[3] M. Mikami, T. Yasuda, A. Terao, M. Nakamura, S. Ueno, H. Tanabe, T. Tanaka, T. Onuma, Y. Goto, S. Kaneko, A. Sano, Localization of a gene for benign adult familial myoclonic epilepsy to chromosome 8q23.3–q24.1, Am. J. Hum. Genet. 65 (1999) 745–751.

[4] C. Charlier, N.A. Singh, S.G. Ryan, T.B. Lewis, B.E. Reus, R.J. Leach, M. Leppert, A pore mutation in a novel KQT-like potassium channel gene in an idiopathic epilepsy family, Nat. Genet. 18 (1998) 53–55.

[5] R. Robinson, N. Taske, T. Sander, A. Heils, W. Whitehouse, F. Goutieres, J. Aicardi, A.E. Lehesjoki, A. Siren, M. Laue Friis, M.J. Kjeldsen, C. Panayiotopoulos, C. Kennedy, C. Ferrie, M. Rees, R.M. Gardiner, Linkage analysis between childhood absence epilepsy and genes encoding GABAA and GABAB receptors, voltage-dependent calcium channels, and the ECA1 region on chromosome 8q, Epilepsy Res. 48 (2002) 169–179.

[6] R.H. Wallace, D.W. Wang, R. Singh, I.E. Scheffer, A.L. George Jr., H.A. Phillips, K. Saar, A. Reis, E.W. Johnson, G.R. Sutherland, S.F. Berkovic, J.C. Mulley, Febrile seizures and generalized epilepsy associated with a mutation in the Na+-channel beta1 subunit gene SCN1B, Nat. Genet. 19 (1998) 366–370.

[7] S. Asakawa, I. Abe, Y. Kudoh, N. Kishi, Y. Wang, R. Kubota, J. Kudoh, K. Kawasaki, S. Minoshima, N. Shimizu, Human BAC library: construction and rapid screening, Gene 191 (1997) 69–79.

[8] S. Asakawa, N. Shimizu, High-fidelity digital hybridization screening, Genomics 49 (1998) 209–217.

[9] D.A. Nickerson, V.O. Tobe, S.L. Taylor, PolyPhred: automating the detection and genotyping of single nucleotide substitutions using fluorescence-based resequencing, Nucleic Acids Res. 25 (1997) 2745–2751.

[10] P.C. Sun, R. Uppaluri, A.P. Schmidt, M.E. Pashia, E.C. Quant, J.B. Sunwoo, S.M. Gollin, S.B. Scholnick, Transcript map of the 8p23 putative tumor suppressor region, Genomics 75 (2001) 17–25.

[11] P. Bork, G. Beckmann, The CUB domain. A widespread module in developmentally regulated proteins, J. Mol. Biol. 231 (1993) 539–545.

[12] D.G. Norman, P.N. Barlow, M. Baron, A.J. Day, R.B. Sim, I.D. Campbell, Three-dimensional structure of a complement control protein module in solution, J. Mol. Biol. 219 (1991) 717–725.

[13] N.A. Singh, C. Charlier, D. Stauffer, B.R. DuPont, R.J. Leach, R. Melis, G.M. Ronen, I. Bjerre, T. Quattlebaum, J.V. Murphy, M.L. McHarg, D. Gagnon, T.O. Rosales, A. Peiffer, V.E. Anderson, M. Leppert, A novel potassium channel gene, KCNQ2, is mutated in an inherited epilepsy of newborns, Nat. Genet. 18 (1998) 25–29.

[14] B.C. Schroeder, C. Kubisch, V. Stein, T.J. Jentsch, Moderate loss of function of cyclic-AMP-modulated KCNQ2/KCNQ3 K+ channels causes epilepsy, Nature 396 (1998) 687–690.

[15] L.A. Pennacchio, A.E. Lehesjoki, N.E. Stone, V.L. Willour, K. Virtaneva, J. Miao, E. D'Amato, L. Ramirez, M. Faham, M. Koskiniemi, J.A. Warrington, R. Norio, A. de la Chapelle, D.R. Cox, R.M. Myers, Mutations in the gene encoding cystatin B in progressive myoclonus epilepsy (EPM1), Science 271 (1996) 1731–1734.

[16] M.H. Kim, J.M. Gunnersen, S.S. Tan, Localized expression of the seizure-related gene SEZ-6 in developing and adult forebrains, Mech. Dev. 118 (2002) 171–174.

[17] T. Kitada, S. Asakawa, N. Hattori, H. Matsumine, Y. Yamamura, S. Minoshima, M. Yokochi, Y. Mizuno, N. Shimizu, Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism, Nature 392 (1998) 605–608.